Measure Theory

Table of contents

Chapter 1
Classes

1.1 Basic concepts about classes

We use Von Neuman's set of axioms to define a set and class theory. We start with two undefined notions: **class** and the **membership relation** $\in$. The objects of our discourse will be classes. Later we will introduce the notion of **set** that is a special kind of class. Every **set** will be a class but not every **class** will be a **set**. Intuitively we understand a class to be a kind of collection and $x \in C$ to mean that x is in the collection C. The elements of a class are themselves considered classes. As a notation shorthand we always use uppercase for classes (if we think of them as a collection) and lowercase for elements of a class (although they are also classes). Further we assume that the normal rules of intuitive logic are known.

Notation 1.1. *We introduce the following rotational terms*

$$\begin{array}{rll} \wedge & \text{meaning} & \text{and} \\ \vee & \text{meaning} & \text{or} \\ \neg & \text{meaning} & \text{not} \\ \Rightarrow & \text{meaning} & \text{implies} \\ \Leftrightarrow & \text{meaning} & \text{is equivalent with} \\ \in & \text{meaning} & \text{is element} \\ \vdash & \text{meaning} & \text{with} \\ \vDash & \text{meaning} & \text{where} \\ \forall & \text{meaning} & \text{forall} \\ \exists & \text{meaning} & \text{there exists} \\ \exists! & \text{meaning} & \text{there exists } a \text{ unique} \end{array}$$

Definition 1.2. *We say that x is a element if there exists a class A such that $x \in A$*

Definition 1.3. *$x \notin A$ is the same as $\neg(x \in A)$*

Definition 1.4. *Let A, B be classes then we say that $A = B$ if and only if $x \in A \Rightarrow x \in B \wedge x \in B \Rightarrow x \in A$*

So two classes are considered equal if they have the same elements. Once we have equality defined for classes we can establish our first axiom concerning classes.

Axiom 1.5. (Axiom of extent) *If* $x = y$ *and* $x \in A \Rightarrow y \in A$.

Definition 1.6. *Let* A *and* B *be classes then* $A \subseteq B$ *iff* $x \in A \Rightarrow x \in B$, *we call* A *a subclass of* B

Definition 1.7. *Let* A *and* B *be classes then* $A \subset B$ *iff* $A \subseteq B$ *and* $A \neq B$, *we call* A *a proper subclass of* B

Theorem 1.8. *For all classes* A, B *and* C *the following hold:*

1. $A = A$
2. $A = B \Rightarrow B = A$
3. $A = B$ *and* $B = C \Rightarrow A = C$
4. $A \subseteq B$ *and* $B \subseteq A \Rightarrow A = B$
5. $A \subseteq B$ *and* $B \subseteq C \Rightarrow A \subseteq C$

Proof.

1. $x \in A \Rightarrow x \in A$ and $x \in A \Rightarrow x \in A$ is obviously true, thus by definition 1.4 we have $A = A$.
2. If $A = B$ then $x \in A \Rightarrow x \in B \wedge x \in B \Rightarrow x \in A$ which is equivalent with $x \in B \Rightarrow x \in A \wedge x \in A \Rightarrow x \in B$ proving $B = A$
3. If $A = B$ and $B = C$ then we have $x \in A \Rightarrow x \in B$ and $x \in B \Rightarrow x \in C$ so that we have $x \in A \Rightarrow x \in C$ (a). Also we have $x \in C \Rightarrow x \in B$ and $x \in B \Rightarrow x \in A$ giving thus $x \in C \Rightarrow x \in A$ (b) so by (a) and (b) we have $A = C$
4. From $A \subseteq B$ we have $x \in A \Rightarrow x \in B$ and from $B \subseteq A$ we have $x \in B \Rightarrow x \in A$ proving $A = B$
5. From $A \subseteq B$ it follows that $x \in A \Rightarrow x \in B$ (a) and from $B \subseteq C$ it follows that $x \in B \Rightarrow x \in C$ (b). From (a) and (b) we have then $x \in A \Rightarrow x \in C$ and thus $A \subseteq B$ □

The following axiom provides us with a useful way of creating new classes. Let $P(x)$ denotes a statement about x which can be expressed in terms of the symbols $\in$, $\wedge$, $\vee$, $\neg$, $\Rightarrow$, $\exists$, $\forall, \vdash, \vDash$ brackets and variables $x, y, z \ldots, A, B, C, \ldots$ then we have. Note that in a rigid and formal theory of sets we have to specify the rules to form predicates and propositions so that we avoid any paradoxes. As our purpose of this text is to provide a global overview of analysis we define propositions and predicates in a informal way.

A famous paradox that we want to avoid in this text is Russel's paradox: Consider the class R of all classes that do not contain them self and ask if R contains itself? If R contains itself then by definition of R it may not contains itself and if it does not contains itself then by definition of R it will contains itself. The solution to this paradox is the axiom of class construction that allows us to construct in a safe way new classes based on a statement.

Axiom 1.9. (Axiom of class construction) *There exists a class C such that $x \in C$ iff x is a element and $P(x)$ is true, we note this as $C = \{x | P(x)\}$*

Note 1.10. The above axiom avoids the classical Russel's Paradox as this allows us to form a class consisting of all elements x which satisfies $P(x)$ but not the class of all classes x such that $P(x)$. So if we take $P(x)$ to be $x \notin x$ then $R = \{x | x \notin x\}$ is Russel's class and we have the following two possibilities

1. $R \in R$ is impossible as this would imply that R is a element and $R \notin R$
2. $R \notin R$ then we have the following possibilities
   a. R is a element but then $R \in R$, a contradiction
   b. R is not a element so that indeed $R \notin R$

So we have reduced Russel's paradox to the harmless statement that $R \notin R$ and R is not a element. The use of lowercase in $C = \{x | P(x)\}$ is on purpose as we have agreed to designate elements by lower cases so that we will remember that $C = \{x : P(x)\}$ is the same as $C = \{x | x \,\text{is}\, a \,\text{element and}\, P(x)\}$. Another notation we use is $\{x \in A | P(x)\}$ which is the same as $\{x | x \in A \wedge P(x)\}$ (note that $x \in A$ already implements that x is a element).

Using the axiom of class construction we can now define the union of two classes

Definition 1.11. *Let A and B be classes then $A \bigcup B = \{x | x \in A \vee x \in B\}$*

Definition 1.12. *Let A and B be classes then $A \bigcap B = \{x | x \in A \wedge x \in B\}$*

Definition 1.13. *We define the universal class $\mathcal{U}$ to be $\mathcal{U} = \{x | x = x\}$ as the class of all elements ($x = x$ is always true).*

Definition 1.14. *We define the empty class $\emptyset$ by $\emptyset = \{x | x \neq x\}$ which is the class with no elements for if $x \in \emptyset$ then $x \neq x$ which is a contradiction.*

Note 1.15. As $\emptyset$ is a empty set we have that $\forall x \in \emptyset$ every proposition about x is true.

Theorem 1.16. *For every class A we have $\emptyset \subseteq A$ and $A \subseteq \mathcal{U}$*

Proof.

1. We prove $\emptyset \subseteq A$ by contra-position (meaning $x \in \emptyset \Rightarrow x \in A$ is equivalent with $x \notin A \Rightarrow x \notin \emptyset$). Now if $x \notin A \Rightarrow x \notin \emptyset$ is true as $\emptyset$ has no elements.
2. If $x \in A$ then x is a element and $x = x$ so $x \in \mathcal{U}$ □

Theorem 1.17. *If A is a a class such that $x \in A$ yields a contradiction then $A = \emptyset$*

Proof. As $\forall x \in A$ we have a contradiction it follows that $\forall x \in A \vDash x \neq x$ proving that $A \subseteq \emptyset$ and as we have proved in the previous theorem that $\emptyset \subseteq A$ we conclude that $A = \emptyset$. □

Definition 1.18. *Two classes A, B are disjoint if $A \bigcap B = \emptyset$ (they have no common elements)*

Definition 1.19. *If A is a class then we define $A^c = \{x: x \notin A\}$*

Theorem 1.20. $(A^c)^c = A$

Proof. If $x \in (A^c)^c$ then $\neg(x \in A^c) \Rightarrow \neg(\neg(x \in A)) \Rightarrow x \in A$ and if $x \in A \Rightarrow \neg(\neg x \in A) \Rightarrow \neg(x \in A^c) \Rightarrow x \in (A^c)^c$ □

Definition 1.21. *If A and B are classes then the class $A \backslash B = A \bigcap B^c$.*

Theorem 1.22. *If A and B are classes then $A \backslash B = \{x \in A | x \notin B\}$*

Proof.

$$\begin{aligned} x \in A \backslash B &\Leftrightarrow x \in A \bigcap B^c \\ &\Leftrightarrow x \in A \wedge x \in B^c \\ &\Leftrightarrow x \in A \wedge x \notin B \\ &\Leftrightarrow x \in \{x \in A | x \notin B\} \end{aligned}$$

□

1.2 Class Operations

Theorem 1.23. *If A, B are classes then we have*

1. *$A \subseteq A \bigcup B$ and $B \subseteq A \bigcup B$*
2. *$A \bigcap B \subseteq A$ and $A \bigcap B \subseteq B$*

Proof.

1. $x \in A \Rightarrow x \in A \vee x \in B \Rightarrow x \in A \bigcup B$, $x \in B \Rightarrow x \in A \vee x \in B \Rightarrow x \in A \bigcup B$
2. $x \in A \bigcap B \Rightarrow x \in A \wedge x \in B \Rightarrow x \in A$ and thus $A \bigcap B \subseteq A$. $x \in A \bigcap B \Rightarrow x \in B$ and thus $A \bigcap B \subseteq B$ □

Theorem 1.24. *If A, B are classes then we have*

1. *$A \subseteq B$ if and only if $A \bigcup B = B$*
2. *$A \subseteq B$ if and only if $A \bigcap B = A$*

Proof.

1.

   ⇒. If $x \in A \bigcup B \Rightarrow x \in A \underset{A \subseteq B}{\Rightarrow} x \in B$ and thus $A \bigcup B \subseteq B$. From the previous theorem we have $B \subseteq A \bigcup B$ so by 1.8 we have $A \bigcup B = B$

⇐. If $A\bigcup B=B$ then $x\in A\Rightarrow x\in A\bigcup B \underset{A\bigcup B=B}{\Rightarrow} x\in B$ and thus $A\subseteq B$

2.

⇒. If $x\in A \underset{A\subseteq B}{\Rightarrow} x\in B\Rightarrow x\in A\wedge x\in B\Rightarrow x\in A\bigcap B$ proving that $A\subseteq A\bigcap B$. From the previous theorem we have $A\bigcap B\subseteq A$ so by 1.8 we have $A\bigcap B=A$

⇐. If $A\bigcap B=A$ we have $x\in A\Rightarrow x\in A\bigcap B\Rightarrow(x\in A\wedge x\in B)\Rightarrow x\in B$ so $A\subseteq B$. □

Theorem 1.25. (Absorption Laws) *If A,B are classes then*

1. $A\bigcup(A\bigcap B)=A$
2. $A\bigcap(A\bigcup B)=A$

Proof.

1. *By 1.23 we have $A\bigcap B\subseteq A$ and thus by the previous theorem we have $A\bigcup(A\bigcap B)=A$*
2. *By 1.23 we have $A\subseteq A\bigcup B$ and thus by the previous theorem we have $A\bigcap(A\bigcup B)=A$* □

Theorem 1.26. (DeMorgan's Law) *For all classes A,B we have*

1. $(A\bigcap B)^c=A^c\bigcup B^c$
2. $(A\bigcup B)^c=A^c\bigcap B^c$

Proof.

1. First

$$\begin{aligned} x\in\left(A\bigcap B\right)^c &\Rightarrow x\notin\left(A\bigcap B\right)\\ &\Rightarrow \neg(x\in A\wedge x\in B)\\ &\Rightarrow x\notin A\vee x\notin B\\ &\Rightarrow x\in A^c\vee x\in B^c\\ &\Rightarrow x\in A^c\bigcup B^c \end{aligned}$$

Second

$$\begin{aligned} x\in A^c\bigcup B^c &\Rightarrow x\in A^c\vee x\in B^c\\ &\Rightarrow x\notin A\vee x\notin B\\ &\Rightarrow \neg(x\in A\wedge x\in B)\\ &\Rightarrow \neg\left(x\in A\bigcap B\right)\\ &\Rightarrow x\in\left(A\bigcap B\right)^c \end{aligned}$$

2. First

$$\begin{aligned} x \in \left(A \bigcup B\right)^c &\Rightarrow x \notin \left(A \bigcup B\right) \\ &\Rightarrow \neg\left(x \in A \bigcup B\right) \\ &\Rightarrow \neg(x \in A \vee x \in B) \\ &\Rightarrow x \notin A \wedge x \notin B \\ &\Rightarrow x \in A^c \wedge x \in B^c \\ &\Rightarrow x \in A^c \bigcap B^c \end{aligned}$$

Second

$$\begin{aligned} x \in A^c \bigcap B^c &\Rightarrow x \in A^c \wedge x \in B^c \\ &\Rightarrow x \notin A \wedge x \notin B \\ &\Rightarrow \neg(x \in A \vee x \in B) \\ &\Rightarrow \neg\left(x \in A \bigcup B\right) \\ &\Rightarrow x \notin \left(A \bigcup B\right) \\ &\Rightarrow x \in \left(A \bigcup B\right)^c \end{aligned}$$

□

Theorem 1.27. *If A, B are cflasses then*

1. $A \bigcup B = B \bigcup A$
2. $A \bigcap B = B \bigcap A$

Proof.

1. *We have*

$$\begin{aligned} x \in A \bigcup B &\Leftrightarrow x \in A \vee x \in B \\ &\Leftrightarrow x \in B \vee x \in A \\ &\Leftrightarrow x \in B \bigcup A \end{aligned}$$

2. *We have*

$$\begin{aligned} x \in A \bigcap B &\Leftrightarrow x \in A \wedge x \in B \\ &\Leftrightarrow x \in B \wedge x \in A \\ &\Leftrightarrow x \in B \bigcap A \end{aligned}$$

□

***Theorem* 1.28.** *For any class A we have*

1. $A \bigcup A = A$
2. $A \bigcap A = A$

Proof. *This follows from $A \subseteq A$ and 1.24.* □

***Theorem* 1.29. (associativity)** *For any class A, B, C we have*

1. $A \bigcup (B \bigcup C) = (A \bigcup B) \bigcup C$

2. $A\bigcap(B\bigcap C)=(A\bigcap B)\bigcap C$

Proof.

1. *We have*

$$\begin{aligned} x\in A\bigcup\left(B\bigcup C\right) &\Leftrightarrow x\in A\vee x\in\left(B\bigcup C\right)\\ &\Leftrightarrow x\in A\vee(x\in B\vee x\in C)\\ &\Leftrightarrow (x\in A\vee x\in B)\vee x\in C\\ &\Leftrightarrow x\in A\bigcup B\vee x\in C\\ &\Leftrightarrow x\in\left(A\bigcup B\right)\bigcup C \end{aligned}$$

2. *We have*

$$\begin{aligned} x\in A\bigcap\left(B\bigcap C\right) &\Leftrightarrow x\in A\wedge x\in B\bigcap C\\ &\Leftrightarrow x\in A\wedge(x\in B\wedge x\in C)\\ &\Leftrightarrow (x\in A\wedge x\in B)\wedge x\in C\\ &\Leftrightarrow x\in A\bigcap B\wedge x\in C\\ &\Leftrightarrow x\in\left(A\bigcap B\right)\bigcap C \end{aligned}$$

□

Theorem 1.30. *For every class we have*

1. $A\bigcup\emptyset=A$
2. $A\bigcap\emptyset=\emptyset$
3. $A\bigcup\mathcal{U}=\mathcal{U}$
4. $A\bigcap\mathcal{U}=A$
5. $\mathcal{U}^c=\emptyset$
6. $\emptyset^c=\mathcal{U}$
7. $A\bigcup A^c=\mathcal{U}$
8. $A\bigcap A^c=\emptyset$

Proof.

1. As $\emptyset\subseteq A$ (see 1.16) it follows from 1.24 that $A\bigcup\emptyset=A$
2. As $\emptyset\subseteq A$ (see 1.16) it follows from 1.24 that $A\bigcap\emptyset=\emptyset$
3. As $A\subseteq\mathcal{U}$ (see 1.16) it follows from 1.24 that $A\bigcup\mathcal{U}=\mathcal{U}$
4. As $A\subseteq\mathcal{U}$ (see 1.16) it follows from 1.24 that $A\bigcap\mathcal{U}=A$

5. First $\emptyset \subseteq \mathcal{U}^c$ (see 1.16). Second

$$\begin{aligned} x \in \mathcal{U}^c &\Rightarrow x \notin \mathcal{U} \\ &\Rightarrow x \neq x \\ &\Rightarrow x \in \emptyset \end{aligned}$$

proving $\mathcal{U}^c \subseteq \emptyset$ and thus $\emptyset = \mathcal{U}^c$

6. First we have $\emptyset^c \subseteq \mathcal{U}$ (see 1.16). Second

$$\begin{aligned} x \in \mathcal{U} &\Rightarrow x = x \\ &\Rightarrow \neg(x \neq x) \\ &\Rightarrow x \notin \emptyset \\ &\Rightarrow x \in \emptyset^c \end{aligned}$$

7. We have

$$\begin{aligned} x \in A \bigcup A^c &\Leftrightarrow x \in A \vee x \in A^c \\ &\Leftrightarrow x \in A \vee x \notin A \,(\text{always true}) \\ &\Leftrightarrow x = x \,(\text{always true}) \\ &\Leftrightarrow x \in \mathcal{U} \end{aligned}$$

8. We have

$$\begin{aligned} x \in A \bigcap A^c &\Leftrightarrow x \in A \wedge x \in A^c \\ &\Leftrightarrow x \in A \wedge x \notin A \,(\text{always false}) \\ &\Leftrightarrow x \neq x \,(\text{always false}) \\ &\Leftrightarrow x \in \emptyset \end{aligned}$$

□

Theorem 1.31. *If A, B, C are classes then we have*

1. $A \bigcap (B \bigcup C) = (A \bigcap B) \bigcup (A \bigcap C)$
2. $A \bigcup (B \bigcap C) = (A \bigcup B) \bigcap (A \bigcup C)$
3. $A \setminus (B \bigcup C) = (A \setminus B) \bigcap (A \setminus C)$
4. $A \setminus (B \bigcup C) = (A \setminus B) \setminus C$
5. $A \setminus (B \bigcap C) = (A \setminus B) \bigcup (A \setminus C)$
6. *If* $B \subseteq A$ *then* $A \setminus (A \setminus B) = B$
7. $A \bigcup B = A \bigcup (B \setminus A)$
8. *If* $A \subseteq B$ *then* $B = B \setminus A \bigcup A$
9. $(A \bigcup B) \setminus C = (A \setminus C) \bigcup (B \setminus C)$
10. $(A \bigcap B) \setminus C = (A \setminus C) \bigcap B = A \bigcap (B \setminus C)$

Proof.

1. *We have*

$$\begin{aligned}
x \in A\bigcap\left(B\bigcup C\right) &\Leftrightarrow x \in A \wedge x \in B\bigcup C \\
&\Leftrightarrow x \in A \wedge (x \in B \vee x \in C) \\
&\Leftrightarrow (x \in A \wedge x \in B) \vee (x \in A \wedge x \in C) \\
&\Leftrightarrow x \in A\bigcap B \vee x \in A\bigcap C \\
&\Leftrightarrow x \in \left(A\bigcap B\right)\bigcup\left(A\bigcap C\right)
\end{aligned}$$

2. *We have*

$$\begin{aligned}
x \in A\bigcup\left(B\bigcap C\right) &\Leftrightarrow x \in A \vee x \in \left(B\bigcap C\right) \\
&\Leftrightarrow x \in A \vee (x \in B \wedge x \in C) \\
&\Leftrightarrow (x \in A \vee x \in B) \wedge (x \in A \vee x \in C) \\
&\Leftrightarrow x \in A\bigcup B \wedge x \in A\bigcup C \\
&\Leftrightarrow x \in \left(A\bigcup B\right)\bigcap\left(A\bigcup C\right)
\end{aligned}$$

3. *We have*

$$\begin{aligned}
A\setminus\left(B\bigcup C\right) &\underset{\text{definition of}}{=} A\bigcap\left(B\bigcup C\right)^c \\
&\underset{1.26}{=} A\bigcap\left(B^c\bigcap C^c\right) \\
&= \left(A\bigcap A\right)\bigcap\left(B^c\bigcap C^c\right) \\
&\underset{1.29}{=} \left(A\bigcap\left(B^c\bigcap C^c\right)\right)\bigcap A \\
&\underset{1.29}{=} \left(\left(A\bigcap B^c\right)\bigcap C^c\right)\bigcap A \\
&\underset{1.29}{=} \left(A\bigcap B^c\right)\bigcap\left(C^c\bigcap A\right) \\
&\underset{1.27}{=} \left(A\bigcap B^c\right)\bigcap\left(A\bigcap C^c\right) \\
&= (A\setminus B)\bigcap(A\setminus C)
\end{aligned}$$

4. *We have*

$$\begin{aligned}
A\setminus\left(B\bigcup C\right) &= A\bigcap\left(B\bigcup C\right)^c \\
&= A\bigcap\left(B^c\bigcap C^c\right) \\
&= \left(A\bigcap B^c\right)\bigcap C^c \\
&= (A\setminus B)\setminus C
\end{aligned}$$

5. *We have*

$$\begin{aligned}
A\setminus\left(B\bigcap C\right) &= A\bigcap\left(B\bigcap C\right)^c \\
&\underset{1.26}{=} A\bigcap\left(B^c\bigcup C^c\right) \\
&\underset{(1)}{=} \left(A\bigcap B^c\right)\bigcup\left(A\bigcap C^c\right) \\
&= (A\setminus B)\bigcup(A\setminus C)
\end{aligned}$$

6. *We have if* $B \subseteq A$

$$\begin{aligned} A\setminus(A\setminus B) &= A\bigcap(A\setminus B)^c \\ &= A\bigcap(A\bigcap B^c)^c \\ &\underset{1.26}{=} A\bigcap(A^c\bigcup(B^c)^c) \\ &\underset{1.20}{=} A\bigcap(A^c\bigcup B) \\ &\underset{1.26}{=} (A\bigcap A^c)\bigcup(A\bigcap B) \\ &\underset{1.30}{=} \emptyset\bigcup(A\bigcap B) \\ &\underset{1.30}{=} A\bigcap B \\ &\underset{B\subseteq A}{=} B \end{aligned}$$

7. $A\bigcup B\setminus A = A\bigcup(B\bigcap A^c) = (A\bigcup B)\bigcap(A\bigcup A^c) \underset{1.30}{=} (A\bigcup B)\bigcap\mathcal{U} \underset{1.30}{=} A\bigcup B$
8. *If* $A\subseteq B$ then by (6) we have $A\bigcup B\setminus A = A\bigcup B \underset{1.24}{=} B$
9. $(A\bigcup B)\setminus C = (A\bigcup B)\bigcap C^c = (A\bigcap C^c)\bigcup(B\bigcap C^c) = (A\setminus C)\bigcup(B\setminus C)$
10. $(A\bigcap B)\setminus C = (A\bigcap B)\bigcap C^c = (A\bigcap C^c)\bigcap B = (A\setminus C)\bigcap B$ *and* $(A\bigcap B)\setminus C = (A\bigcap B)\bigcap C^c = A\bigcap(B\bigcap C^c) = A\bigcap(B\setminus C)$

□

Theorem 1.32. *Let* A, B, C *be classes such that* $A\bigcap B = \emptyset$, $C = A\bigcup B$ *then*

1. $A = C\setminus B$
2. $B = C\setminus A$
3. $A\setminus B = A$

Proof.

1. If $x\in A\subseteq A\bigcup B = C \Rightarrow x\in C$ and if $x\in B$ we would have $x\in A\bigcap B = \emptyset$ a contradiction so $x\notin B\Rightarrow x\in C\setminus B\Rightarrow A\subseteq C\setminus B$. If $x\in C\setminus B$ then $x\in C = A\bigcup B$ so $x\in A\vee x\in B$, as $x\notin B$ we must have $x\in A\Rightarrow C\setminus B\subseteq A$.
2. If $x\in B\subseteq A\bigcup B = C\Rightarrow x\in C$ and if $x\in A$ we would have $x\in A\bigcap B = \emptyset$ a contradiction so $x\notin A\Rightarrow x\in C\setminus A\Rightarrow B\subseteq C\setminus A$. If $x\in C\setminus A$ then $x\in C = A\bigcup B$ so $x\in A\vee x\in B$, as $x\notin A$ we must have $x\in B\Rightarrow C\setminus A\subseteq B$
3. $x\in A$ then as $A\bigcap B = \emptyset$ w emust have $x\notin B$ so that $x\in A\setminus B$ hence $A\subseteq A\setminus B\subseteq A$ so that $A\setminus B = A$ □

1.3 Cartesian Products

If a is a element we can use the axiom of class construction 1.9 to define the class $\{x|x = a\}$, this leads to the following definition of a singleton.

Definition 1.33. *If a is a element then $\{a\} = \{x|x = a\}$ is a class containing only one element. The class $\{a\}$ is called a* ***singleton****.*

If a, b are elements then we can define the class $\{x|x = a\vee x = b\}$ consisting of two elements. This leads to the following definition.

Definition 1.34. *If a, b are elements then $\{a, b\} = \{x | x = a \vee x = b\}$ is called a* ***unordered pair****.*

We want to form classes in which $\{a, b\}$ are members, to be able to this, we introduce the following axiom.

Axiom 1.35. (Axiom of Pairing) *If a, b are elements then $\{a, b\}$ is a element*

Lemma 1.36. *If a is a element then $\{a, a\} = \{a\}$*

Proof.

$$\begin{aligned} x \in \{a, a\} &\Leftrightarrow x = a \vee x = a \\ &\Leftrightarrow x = a \\ &\Leftrightarrow x \in \{a\} \end{aligned}$$

□

Theorem 1.37. *If a is a element then $\{a\}$ is a element*

Proof. As a is a element we have that a, a are elements and by the previous axiom we have that $\{a, a\} = \{a\}$ is a element. □

Theorem 1.38. *If x, y, x', y' are elements then from $\{x, y\} = \{x', y'\}$ we have $(x = x' \wedge y = y') \vee (x = y' \wedge y = x')$*

Proof. Lets's consider the following possible cases:

- $\boldsymbol{x = y}$**.** so $\{x, y\} = \{x\} = \{x', y'\}$ and thus as $x' \in \{x', y'\} = \{x\}$ we must have $x = x'$ and $y' \in \{x', y'\} = \{x\} \Rightarrow y' = x = y$ proving that $(x = x' \wedge y = y') \Rightarrow (x = x' \wedge y = y') \vee (x = y' \wedge y = x')$
- $\boldsymbol{x \neq y}$**.** then we have from $\{x, y\} = \{x', y'\}$ the following case to consider
  - $\boldsymbol{x = x'}$**.** then as $y \in \{x', y'\}$ and $x = x' \neq y$ we must have $y = y'$ proving that $x = x' \wedge y = y' \Rightarrow (x = x' \wedge y = y') \vee (x = y' \wedge y = x')$
  - $\boldsymbol{x = y'}$**.** then as $y \in \{x', y'\}$ and $y' = x \neq y$ we must have $y = x'$ proving that $x = y' \wedge y = x' \Rightarrow (x = x' \wedge y = y') \vee (x = y' \wedge y = x')$

  So we have always $(x = x' \wedge y = y') \vee (x = y' \wedge y = x')$ □

We can now define a ordered pair of elements

Definition 1.39. *If a, b are elements then $(a, b) = \{\{a\}, \{a, b\}\}$ (which is a element as $\{a, b\}$ and $\{a\}$ are elements)*

Theorem 1.40. *If $(a, b) = (c, d)$ then $a = c$ and $b = d$*

Proof. If $(a, b) = (c, d)$ then we have by definition $\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}$. Using 1.38 we have then the following possible cases

- $\boldsymbol{\{a\} = \{c\} \wedge \{a, b\} = \{c, d\}}$**.** from this we conclude that $a = c$ and using 1.38 again we have either
  - $\boldsymbol{a = c \wedge b = d}$**.** giving the required result

$a=d\wedge b=c$. then as $a=c$ we have $c=a=d\wedge b=d\Rightarrow a=c\wedge b=d$

$\{a\}=\{c,d\}\wedge\{a,b\}=\{c\}$. then we have $a=c=d\wedge a=b=c\Rightarrow a=c\wedge b=d$ □

Definition 1.41. *If A,B are classes then the class $A\times B=\{z|z=(a,b)\text{ where }a\in A\wedge b\in B\}$ is noted as $A\times B=\{(x,y)|x\in A\wedge y\in B\}$ and is named the **Cartesian product** of A and B (note that $A\times B$ is a class as $x\in A\wedge y\in B$ means that x,y are elements so that (x,y) is a element)*

Example 1.42. $\emptyset=\emptyset\times\emptyset$

Proof. If $z\in\emptyset\times\emptyset$ then there exists $x,y\in\emptyset$ such that $z=(x,y)$ which is a contradiction so by 1.17 we have that $\emptyset\times\emptyset=\emptyset$. □

Actually we can extend the above example to a more general case

Lemma 1.43. *If A,B are classes then we have $A\times B=\emptyset\Leftrightarrow A=\emptyset\vee B=\emptyset$*

Proof.

$\Rightarrow$. Let $A\times B=\emptyset$ and assume that $A\neq\emptyset\wedge B\neq\emptyset$ then $\exists x\in A\wedge\exists y\in B$ such that $(x,y)\in A\times B\Rightarrow A\times B\neq\emptyset$ a contradiction. So $A=\emptyset\vee B=\emptyset$

$\Leftarrow$. Let $A=\emptyset\vee B=\emptyset$ and assume that $A\times B\neq\emptyset$ then there exists a $z\in A\times B$ such that $z=(x,y)$ where $x\in A\wedge y\in B$ giving $A\neq\emptyset\wedge B=\emptyset$ contradicting $A=\emptyset\vee B=\emptyset$. So $A\times B=\emptyset$. □

The following is a consequence of this definition

Lemma 1.44. *If A,B,C,D are classes with $A\subseteq B$ and $C\subseteq D$ then $A\times C\subseteq B\times D$*

Proof.

$$\begin{aligned} z\in A\times C &\Rightarrow z=(x,y)\wedge x\in A\wedge x\in C\\ &\Rightarrow z=(x,y)\wedge x\in B\wedge x\in D\\ &\Rightarrow z\in B\times D \end{aligned}$$

Proving that $A\times C\subseteq B\times D$ □

For non empty sets we can have the opposite relations

Lemma 1.45. *If A,B,C,D are classes with $A\neq\emptyset\wedge B\neq\emptyset$ then if $A\times B\subseteq C\times D$ we have $A\subseteq C\wedge B\subseteq D$*

Proof. If $x\in A$ then as $B\neq\emptyset$ there exists a $y\in B$ so that $(x,y)\in A\times B\subseteq C\times D$ hence $(x,y)\in C\times D$ proving that $x\in C$, so $A\subseteq C$. Further if $y\in B$ then as $A\neq\emptyset$ there exists a $x\in A$ so that $(x,y)\in A\times B\subseteq C\times D$ hence $(x,y)\in C\times D$ proving that $y\in D$, so $B\subseteq D$. □

Combining the Cartesian with the other class operations give

Lemma 1.46. *For all classes A, B, C and D we have*

1. $A \times (B \bigcap C) = (A \times B) \bigcap (A \times C)$
2. $A \times (B \bigcup C) = (A \times B) \bigcup (A \times C)$
3. $(A \times B) \bigcap (C \times D) = (A \bigcap C) \times (B \bigcap D)$
4. $(B \bigcap C) \times A = (B \times A) \bigcap (C \times A)$
5. $(B \bigcup C) \times A = (B \times A) \bigcup (C \times A)$
6. $(A \times B) \setminus (C \times D) = ((A \setminus C) \times B) \bigcup (A \times (B \setminus D))$
7. $(A \setminus B) \times C = (A \times C) \setminus (B \times C)$
8. $A \times (B \setminus C) = (A \times B) \setminus (A \times C)$

Proof.

1. *We have*

$$\begin{aligned} z \in A \times \left(B \bigcap C\right) &\Leftrightarrow z = (x, y) \wedge x \in A \wedge y \in \left(B \bigcap C\right) \\ &\Leftrightarrow z = (x, y) \wedge x \in A \wedge (y \in B \wedge y \in C) \\ &\Leftrightarrow (z = (x, y) \wedge x \in A \wedge y \in B) \wedge (z = (x, y) \wedge x \in A \wedge y \in C) \\ &\Leftrightarrow z \in A \times B \wedge z \in A \times C \\ &\Leftrightarrow z \in (A \times B) \bigcap (A \times C) \end{aligned}$$

2. *We have*

$$\begin{aligned} z \in A \times \left(B \bigcup C\right) &\Leftrightarrow z = (x, y) \wedge x \in A \wedge y \in \left(B \bigcup C\right) \\ &\Leftrightarrow z = (x, y) \wedge x \in A \wedge (y \in B \vee y \in C) \\ &\Leftrightarrow (z = (x, y) \wedge x \in A \wedge y \in B) \vee (z = (x, y) \wedge x \in A \wedge y \in C) \\ &\Leftrightarrow z \in A \times B \vee z \in A \times C \\ &\Leftrightarrow z \in (A \times B) \bigcup (A \times C) \end{aligned}$$

3. *We have*

$$\begin{aligned} z \in (A \times B) \bigcap (C \times D) &\Leftrightarrow z \in A \times B \wedge z \in C \times D \\ &\Leftrightarrow (z = (x, y) \wedge x \in A \wedge y \in B) \wedge (z = (x', y') \wedge x' \in C \wedge y' \in D) \\ &\underset{(x,y)=z=(x',y') \Rightarrow x=x', y=y'}{\Leftrightarrow} z = (x, y) \wedge x \in A \wedge y \in B \wedge x \in C \wedge y \in D \\ &\Leftrightarrow z = (x, y) \wedge (x \in A \wedge x \in C) \wedge (y \in B \wedge y \in D) \\ &\Leftrightarrow z = (x, y) \wedge \left(x \in A \bigcap C\right) \wedge \left(y \in B \bigcap D\right) \\ &\Leftrightarrow z \in \left(A \bigcap C\right) \times \left(B \bigcap D\right) \end{aligned}$$

4. *We have*

$$\begin{aligned} z\in \big(B\bigcap C\big)\times A &\Leftrightarrow z=(x,y)\wedge x\in B\bigcap C\wedge y\in A \\ &\Leftrightarrow z=(x,y)\wedge x\in B\wedge x\in C\wedge y\in A \\ &\Leftrightarrow (z=(x,y)\wedge x\in B\wedge y\in A)\wedge(z=(x,y)\wedge x\in C\wedge y\in A) \\ &\Leftrightarrow z\in B\times A\wedge z\in C\times A \\ &\Leftrightarrow z\in(B\times A)\bigcap(C\times A) \end{aligned}$$

5. *We have*

$$\begin{aligned} z\in \big(B\bigcup C\big)\times A &\Leftrightarrow z=(x,y)\wedge x\in B\bigcup C\wedge y\in A \\ &\Leftrightarrow z=(x,y)\wedge(x\in B\vee x\in C)\wedge y\in A \\ &\Leftrightarrow (z=(x,y)\wedge x\in B\wedge y\in A)\vee(z=(x,y)\wedge x\in C\wedge y\in A) \\ &\Leftrightarrow (z\in B\times A)\vee(z\in C\times A) \\ &\Leftrightarrow z\in(B\times A)\bigcup(C\times A) \end{aligned}$$

6. *We have*

$$\begin{aligned} z\in(A\times B)\setminus(C\times D) &\Leftrightarrow z\in A\times B\wedge z\notin C\times D \\ &\Leftrightarrow z=(x,y)\wedge(x,y)\in A\times B\wedge(x,y)\notin C\times D \\ &\Leftrightarrow z=(x,y)\wedge x\in A\wedge y\in B\wedge\neg(x\in C\wedge y\in D) \\ &\Leftrightarrow z=(x,y)\wedge x\in A\wedge y\in B\wedge(x\notin C\vee y\notin D) \\ &\Leftrightarrow (z=(x,y)\wedge x\in A\wedge y\in B\wedge x\notin C)\vee(z=(x,y)\wedge \\ &\qquad x\in A\wedge y\in B\wedge y\notin D) \\ &\Leftrightarrow (z\in(A\setminus C)\times B)\vee(z\in A\times(B\setminus D)) \\ &\Leftrightarrow z\in((A\setminus C)\times B)\bigcup(A\times(B\setminus D)) \end{aligned}$$

7. *We have*

$$\begin{aligned} z\in(A\setminus B)\times C &\Rightarrow z=(x,y)\wedge x\in A\setminus B\wedge y\in C \\ &\Rightarrow z=(x,y)\wedge x\in A\wedge x\notin B\wedge y\in C \\ &\Rightarrow (z=(x,y)\wedge x\in A\wedge y\in C)\wedge(z=(x,y)\wedge x\notin B\wedge y\in C) \\ &\Rightarrow z\in A\times C\wedge z\notin B\times C \\ &\Rightarrow z\in(A\times C)\setminus(B\times C) \end{aligned}$$

and

$$\begin{aligned} z\in(A\times C)\setminus(B\times C) &\Rightarrow z=(x,y)\wedge[(x,y)\in A\times C\wedge(x,y)\notin(B\times C)] \\ &\Rightarrow z=(x,y)\wedge[(x\in A\wedge y\in C)\wedge(x\notin B\vee y\notin C)] \\ &\Rightarrow z=(x,y)\wedge[(x\in A\wedge y\in C\wedge x\notin B)\vee(x\in A\wedge \\ &\qquad y\in C\wedge y\notin C)] \\ &\Rightarrow z=(x,y)\wedge(x\in A\wedge y\in C\wedge x\notin B) \\ &\Rightarrow z=(x,y)\wedge(x,y)\in(A\setminus B)\times C \\ &\Rightarrow z\in(A\setminus B)\times C \end{aligned}$$

8. *We have*

$$\begin{aligned} z\in A\times(B\setminus C) &\Rightarrow z=(x,y)\wedge x\in A\wedge y\in B\setminus C\\ &\Rightarrow z=(x,y)\wedge x\in A\wedge y\in B\wedge y\notin C\\ &\Rightarrow (z=(x,y)\wedge x\in A\wedge y\in B)\wedge(z=(x,y)\wedge x\in A\wedge y\notin C)\\ &\Rightarrow z\in A\times B\wedge z\notin A\times C\\ &\Rightarrow z\in(A\times B)\setminus(A\times C)\end{aligned}$$

and

$$\begin{aligned} z\in(A\times B)\setminus(A\times C) &\Rightarrow z=(x,y)\wedge[(x,y)\in A\times B\wedge(x,y)\notin(A\times C)]\\ &\Rightarrow z=(x,y)\wedge[(x\in A\wedge y\in B)\wedge(x\notin A\vee y\notin C)]\\ &\Rightarrow z=(x,y)\wedge[(x\in A\wedge y\in B\wedge x\notin A)\vee(x\in A\wedge\\ &\quad\ y\in B\wedge y\notin C)]\\ &\Rightarrow z=(x,y)\wedge(x\in A\wedge y\in B\wedge y\notin C)\\ &\Rightarrow z=(x,y)\wedge(x,y)\in A\times(B\setminus C)\\ &\Rightarrow z\in A\times(B\setminus C)\end{aligned}$$

□

Notation 1.47. *From now on we note the class $\{z|z=(x,y)\wedge P(x,y)\}$ as $\{(x,y)|P(x,y)\}$. This makes sense for if we want to prove that a class A of tuples is equal to a class B of tuples then we must prove that $z\in A=\{(x,y)|P(x,y)\}\Leftrightarrow z\in B=\{(x,y)|Q(x,y)\}$ but this is equivalent to $z=(x,y)\wedge P(x,y)\Leftrightarrow z=(x',y')\wedge Q(x',y')$ which because of the fact that $(x,y)=z=(x',y')$ implies that $x=x'\wedge y=y'$ is equivalent with $P(x,y)=Q(x,y)$.*

1.4 Graphs

To define relations, partial functions, families we need the concept of graphs.

Definition 1.48. *A graph is a subclass of $\mathcal{U}\times\mathcal{U}$ or in other words a graph is a class of ordered pairs*

Example 1.49. If A,B are classes then $A\times B$ is a class (see 1.41), as $A,B\subseteq\mathcal{U}$ we have by 1.44 that $A\times B\subseteq\mathcal{U}\times\mathcal{U}$ and thus $A\times B$ is a graph.

Lemma 1.50. *If G,H are graphs then $G\bigcup H$ is a graph*

Proof. As G,H are graphs we have G,H are classes and $G\subseteq\mathcal{U}\times\mathcal{U}\wedge H\subseteq\mathcal{U}\times\mathcal{U}\Rightarrow G\bigcup H\subseteq\mathcal{U}\times\mathcal{U}$ proving that the class $G\bigcup H$ is a graph. □

Definition 1.51. *If G is a graph then $G^{-1}=\{(x,y)|(y,x)\in G\}$*

Lemma 1.52. *If G is a graph then G^{-1} is a graph*

Proof. If $(x,y)\in G^{-1}$ then $(y,x)\in G\subseteq\mathcal{U}\times\mathcal{U}\Rightarrow y\in\mathcal{U}\wedge x\in\mathcal{U}\Rightarrow(x,y)\in\mathcal{U}\times\mathcal{U}$ □

Definition 1.53. *If G,H are graphs then $G\circ H=\{(x,y)|\exists z\vdash(x,z)\in H\wedge(z,x)\in G\}\subseteq\mathcal{U}\times\mathcal{U}$ is a graph, the composition of two graphs.*

Theorem 1.54. *If G, H, J are graphs then we have*

1. $(G \circ H) \circ J = G \circ (H \circ J)$
2. $(G^{-1})^{-1} = G$
3. $(G \circ H)^{-1} = H^{-1} \circ G^{-1}$

Proof.

1. We have

$$\begin{aligned}(x, y) \in (G \circ H) \circ J &\Leftrightarrow \exists z \vdash (x, z) \in J \wedge (z, y) \in (G \circ H)\\ &\Leftrightarrow \exists z, \exists z' \vdash (x, z) \in J \wedge (z, z') \in H \wedge (z', y) \in G\\ &\Leftrightarrow \exists z' \vdash (x, z') \in H \circ J \wedge (z', y) \in G\\ &\Leftrightarrow (x, y) \in G \circ (H \circ J)\end{aligned}$$

2. We have

$$\begin{aligned}(x, y) \in (G^{-1})^{-1} &\Leftrightarrow (y, x) \in G^{-1}\\ &\Leftrightarrow (x, y) \in G\end{aligned}$$

3. We have

$$\begin{aligned}(x, y) \in (G \circ H)^{-1} &\Leftrightarrow (y, x) \in G \circ H\\ &\Leftrightarrow \exists z \vdash (y, z) \in H \wedge (z, x) \in G\\ &\Leftrightarrow \exists z \vdash (z, y) \in H^{-1} \wedge (x, z) \in G^{-1}\\ &\Leftrightarrow (x, y) \in H^{-1} \circ G^{-1}\end{aligned}$$

□

Definition 1.55. *If G is a graph then* $\operatorname{dom}(G)$ *is the class defined by* $\operatorname{dom}(G) = \{x | \exists y \vdash (x, y) \in G\}$

Definition 1.56. *If G is a graph then* $\operatorname{range}(G)$ *is the class defined by* $\operatorname{range}(G) = \{y | \exists x \vdash (x, y) \in G\}$

Theorem 1.57. *If G and H are graphs then*

1. $\operatorname{dom}(G) = \operatorname{range}(G^{-1})$
2. $\operatorname{dom}(G \circ H) \subseteq \operatorname{dom}(H)$
3. $\operatorname{range}(G) = \operatorname{dom}(G^{-1})$
4. $\operatorname{range}(G \circ H) \subseteq \operatorname{range}(G)$

Proof.

1.

$$\begin{aligned}x \in \operatorname{dom}(G) &\Leftrightarrow \exists y \vdash (x, y) \in G\\ &\Leftrightarrow \exists y \vdash (y, x) \in G^{-1}\\ &\Leftrightarrow x \in \operatorname{range}(G^{-1})\end{aligned}$$

2.

$$\begin{aligned}x \in \operatorname{dom}(G \circ H) &\Rightarrow \exists y \vdash (x, y) \in G \circ H\\ &\Rightarrow \exists z \vdash (x, z) \in H \wedge (z, y) \in G\\ &\Rightarrow x \in \operatorname{dom}(H)\end{aligned}$$

3.

$$\begin{aligned} x \in \operatorname{range}(G) &\Leftrightarrow \exists y \vdash (y, x) \in G \\ &\Leftrightarrow \exists y \vdash (x, y) \in G^{-1} \\ &\Leftrightarrow x \in \operatorname{dom}(G^{-1}) \end{aligned}$$

4.

$$\begin{aligned} x \in \operatorname{range}(G \circ H) &\Rightarrow \exists y \vdash (y, x) \in G \circ H \\ &\Rightarrow \exists z \vdash (y, z) \in H \wedge (z, x) \in G \\ &\Rightarrow x \in \operatorname{range}(G) \end{aligned}$$

□

Theorem 1.58. *If* G, H *are graphs with* $\operatorname{range}(H) \subseteq \operatorname{dom}(G)$ *then* $\operatorname{dom}(G \circ H) = \operatorname{dom}(H)$

Proof. Using the previous theorem we have $\operatorname{dom}(G \circ H) \subseteq \operatorname{dom}(H)$

$$\begin{aligned} x \in \operatorname{dom}(H) &\Rightarrow \exists y \vdash (x, y) \in H \\ &\Rightarrow y \in \operatorname{range}(H) \\ &\underset{\operatorname{rang}(H) \subseteq \operatorname{dom}(G)}{\Rightarrow} y \in \operatorname{dom}(G) \\ &\Rightarrow \exists z \vdash (y, z) \in G \\ &\Rightarrow (x, z) \in G \\ &\Rightarrow x \in \operatorname{dom}(G \circ H) \end{aligned}$$

hence $\operatorname{dom}(H) \subseteq \operatorname{dom}(G \circ H)$ proving $\operatorname{dom}(G \circ H) = \operatorname{dom}(H)$. □

Theorem 1.59. *If* G *and* H *are graphs then*

1. $\operatorname{dom}(G \bigcup H) = \operatorname{dom}(G) \bigcup \operatorname{dom}(H)$
2. $\operatorname{range}(G \bigcup H) = \operatorname{range}(G) \bigcup \operatorname{range}(H)$

Proof.

1.

$$\begin{aligned} x \in \operatorname{dom}\left(G \bigcup H\right) &\Leftrightarrow \exists y \vdash (x, y) \in G \bigcup H \\ &\Leftrightarrow \exists y \vdash (x, y) \in G \vee (x, y) \in H \\ &\Leftrightarrow (\exists y \vdash (x, y) \in G) \vee (\exists y \vdash (x, y) \in H) \\ &\Leftrightarrow x \in \operatorname{dom}(G) \bigcup \operatorname{dom}(H) \end{aligned}$$

2.

$$\begin{aligned} y \in \operatorname{range}\left(G \bigcup H\right) &\Leftrightarrow \exists x \vdash (x, y) \in G \bigcup H \\ &\Leftrightarrow \exists x \vdash (x, y) \in G \vee (x, y) \in H \\ &\Leftrightarrow (\exists x \vdash (x, y) \in G) \vee (\exists x \vdash (x, y) \in H) \\ &\Leftrightarrow x \in \operatorname{range}(G) \bigcup \operatorname{range}(H) \end{aligned}$$

□

A following kind of graph will be used to define functions.

Definition 1.60. (function graph) *A graph G is a function graph if* $\forall(x,y),(x,y')\in G$ *we have* $y=y'$

Theorem 1.61. *If F,G are function graphs then $G\circ F$ is a function graph*

Proof. If $(x,z),(x,z')$ then there exists y,y' such that $(x,y),(x,y')\in F$ and $(y,z),(y',z')\in G$. As F is a function graph we have that $y=y'$. So $(y,z),(y,z')\in F$ then $z=z'$ as G is function graph, proving that $G\circ F$ is a function graph. □

1.5 Sets

A special type of classes are classes that them self are elements, we call this classes sets. The concept that not all classes are sets is essential to avoid for example Russel's paradox as explained earlier. We also need some extra axiom's to complete the foundational base of set theory.

Definition 1.62. *A set is a class that is an element of a class. In formula form* x *is a set* $\Leftrightarrow\exists\mathcal{A}$ *with* $x\in\mathcal{A}$

Rephrasing the axiom of pairing we have then

Theorem 1.63. *If x* and *y are sets then $\{x,y\}$ is a set*

Axiom 1.64. (Axiom of Subsets) *Every subclass of a set is a set*

Theorem 1.65. *If A is a set and B a class then $A\bigcap B$ is a set. In particular we have that the intersection of two sets is a set*

Proof. By theorem 1.23 we have $A\bigcap B\subseteq A$ and thus by the axiom of subsets (see 1.64) we have $A\bigcap B$ is a subset □

Axiom 1.66. (Axiom of Unions) *If $\mathcal{A}$ is a set of sets then $\bigcup_{A\in\mathcal{A}}A$ is a set where* $\bigcup_{A\in\mathcal{A}}A=\{x|\exists A\in\mathcal{A}$ *such that* $x\in A\}$

Theorem 1.67. *If A,B are sets then $A\bigcup B$ is a set*

Proof. Take $\mathcal{A}=\{A,B\}$ then $\mathcal{A}$ is a set by 1.63 and thus a set of sets. Using the axiom of unions we have then that $\bigcup_{C\in\mathcal{A}}C$ is a set. Next we prove that $A\bigcup B=\bigcup_{C\in\mathcal{A}}C$

$$\begin{aligned} x\in A\bigcup B &\Leftrightarrow x\in A\vee x\in B\\ &\Leftrightarrow \exists C\in\{A,B\}\vdash x\in C\\ &\Leftrightarrow x\in\bigcup_{C\in\mathcal{A}}C \end{aligned}$$

□

Now if A is a set then if $B \subseteq A$ we have that B is a set so $B \subseteq A$ is equivalent with B is a element and $B \subseteq A$. Using 1.9 we have then that $\{B|B \subseteq A\}$ is a class (as $B \subseteq A$ is equivalent with B is a set and $B \subseteq A$). This leads to the following definition.

Definition 1.68. *If A is a set then $\mathcal{P}(A)$ is the class defined by $\mathcal{P}(A) = \{B|B \subseteq A\}$*

We state now that $\mathcal{P}(A)$ is a set

Axiom 1.69. (Axiom of Power Sets) *If A is a set then $\mathcal{P}(A)$ is a set*

Theorem 1.70. *If A is a set and P a predicate about X then $\{X|X \subseteq A \wedge P(X)\}$ is a set*

Proof. First as $X \subseteq A \Rightarrow X$ is a set and thus a element (see 1.64) so we have that $X \subseteq A \wedge P(X)$ is equivalent with $(X \text{ is } a \text{ element} \wedge X \subseteq A \wedge P(X))$ and thus $B = \{X|X \subseteq A \wedge P(X)\}$ is a class. Now if $X \in B \Rightarrow X \subseteq A \Rightarrow X \in \mathcal{P}(A) \Rightarrow B \subseteq \mathcal{P}(X)$ and thus by the axiom of subsets 1.64 we have that B is a set. □

Not every class is a set as the following theorem proves:

Theorem 1.71. *The universal class $\mathcal{U} = \{x|x \text{ is } a \text{ element and } x = x\}$ is not a set*

Proof. If $\mathcal{U}$ is a set (thus a element) then $\mathcal{U} \in \mathcal{U}$. Define now the class $\mathcal{B} = \{x|x \text{ is } a \text{ element and } x \notin x\}$ then if $x \in \mathcal{B}$ we have x is a element and thus $x \in \mathcal{U}$ so we have $\mathcal{B} \subseteq \mathcal{U}$ and thus by the axiom of subsets (see 1.64) we must have that $\mathcal{B}$ is a set. If now $\mathcal{B} \in \mathcal{B}$ then we have as $\mathcal{B}$ is a set that $\mathcal{B} \notin \mathcal{B}$ contradicting $\mathcal{B} \in \mathcal{B}$ so we must have that $\mathcal{B} \notin \mathcal{B}$ but this leads then to $\mathcal{B} \in \mathcal{B}$ again a contradiction, so in all cases we have a contradiction and thus our initial assumption that $\mathcal{U}$ is a set must be false. □

Definition 1.72. *Let A be a set then $\mathcal{P}'(A) = \mathcal{P}(A) \backslash \{\emptyset\}$ (this is a set as $\mathcal{P}(A)$ is a set (by1.69 and 1.64). It is essentially the set of all nonempty subsets of A*

If A is a set then $\{A\} = \{A, A\}$ is a set by 1.63 and thus by 1.67 $A \bigcup \{A\}$ is a set. This leads to the following definition

Definition 1.73. *If A is a set then the set $s(A)$ is defined by $s(A) = A \bigcup \{A\}$*

Definition 1.74. (Successor Set) *We say that a set A is a successor set iff*

1. $\emptyset \in A$
2. *If* $X \in A \Rightarrow s(X) \in A$

We have now the axiom of infinity, which will be used later to define the set of natural numbers, from which the whole numbers and finally the reals will follow. It is called the axiom of infinity because it leads to the existence of sets with infinite number of members.

Axiom 1.75. (Axiom of Infinity) *There exists a successor set*

One consequence of this theorem is that $\emptyset$ is a set

Theorem 1.76. *$\emptyset$ is a set*

Proof. By the axiom of infinity there exists a successor set A. But then $\emptyset \in A$ and thus the class $\emptyset$ is a set. $\square$

Example 1.77. The set $0 = \emptyset$ it has no elements, the set 1 is the set $\{\emptyset\} = \{0\}$ it has one element, and 2 is the set $\{\emptyset, \{\emptyset\}\} = \{0, 1\}$ which has two elements (for if $\emptyset = \{\emptyset\}$ we would have a contradiction from $\emptyset \in \{\emptyset\} \Rightarrow \emptyset \in \emptyset \underset{\emptyset = \{x | x \neq x |\}}{\Rightarrow} \emptyset \neq \emptyset$)

Let's now proceed to prove that the product of sets is a set, first we need the following lemma.

Lemma 1.78. *If A, B are sets then $A \times B \subseteq \mathcal{P}(\mathcal{P}(A \bigcup B))$*

Proof. We have

$$\begin{aligned}
(x, y) \in A \times B &\Rightarrow (x, y) = \{\{x\}, \{x, y\}\} \wedge x \in A \wedge x \in B \\
&\Rightarrow \{x\} \subseteq A \subseteq A \bigcup B \wedge \{x, y\} \subseteq A \bigcup B \\
&\Rightarrow \{x\} \subseteq A \bigcup B \wedge \{x, y\} \subseteq A \bigcup B \\
&\Rightarrow \{x\} \in \mathcal{P}\left(A \bigcup B\right) \wedge \{x, y\} \in \mathcal{P}\left(A \bigcup B\right) \\
&\Rightarrow \{\{x\}, \{x, y\}\} \subseteq \mathcal{P}\left(A \bigcup B\right) \\
&\Rightarrow \{\{x\}, \{x, y\}\} \in \mathcal{P}\left(\mathcal{P}\left(A \bigcup B\right)\right) \\
&\Rightarrow (x, y) \in \mathcal{P}\left(\mathcal{P}\left(A \bigcup B\right)\right)
\end{aligned}$$

$\square$

Theorem 1.79. *If A, B are sets then $A \times B$ is a set*

Proof. By 1.67 we have that $A \bigcup B$ is a set, so that by 1.69 $\mathcal{P}(A \bigcup B)$ is set and again by 1.69 we have that $\mathcal{P}(\mathcal{P}(A \bigcup B))$ is a set. Now by the previous lemma we have that $A \times B \subseteq \mathcal{P}(\mathcal{P}(A \bigcup B))$ and using 1.64 we have then that $A \times B$ is a set. $\square$

1.6 Pair and triples of classes

Although we have already defined the concept of a pair, we can not simple extend this to pairs (and later triples of classes). If A, B are pure classes (classes that are not elements) then we can not just form $(A, B) = \{A, \{B\}\}$ because this would mean that A, B are elements and not pure classes. So we need another way of forming pairs, triples and so on.

Definition 1.80. *A pair $\langle A, B \rangle$ of classes is the class defined by $(A \times \{\emptyset\}) \bigcup \{B \times \{\{\emptyset\}\}\}$ [which is a class because $\emptyset$ is a set (see 1.76) so $\{\emptyset\}$* and *$\{\{\emptyset\}\}$ are sets (see 1.63) and thus $\emptyset, \{\emptyset\}, \{\{\emptyset\}\}$ are classes which means that $x \times \{\emptyset\}, \{y \times \{\{\emptyset\}\}\}$ are defined and thus their union is a class].*

Theorem 1.81. *If* $\langle A, B\rangle$ *and* $\langle A', B'\rangle$ *are pairs then* $\langle A, B\rangle = \langle A', B'\rangle$ *iff* $A = A' \wedge B = B'$

Proof.

⇒. Assume $\langle A, B\rangle = \langle A', B'\rangle$ then we have

$$\begin{aligned} e \in A &\Rightarrow (e, \emptyset) \in A \times \{\emptyset\} \\ &\Rightarrow (e, \emptyset) \in \langle A, B\rangle \\ &\Rightarrow (e, \emptyset) \in \langle A', B'\rangle \\ &\Rightarrow (e, \emptyset) \in A' \times \{\emptyset\} \text{ (as } (e, \emptyset) \notin B' \times \{\{\emptyset\}\}) \\ &\Rightarrow e \in A' \\ e \in A' &\Rightarrow (e, \emptyset) \in A' \times \{\emptyset\} \\ &\Rightarrow (e, \emptyset) \in \langle A', B'\rangle \\ &\Rightarrow (e, \emptyset) \in \langle A, B\rangle \\ &\Rightarrow (e, \emptyset) \in A \times \{\emptyset\} \text{ as } (e, \emptyset) \notin B \times \{\{\emptyset\}\} \\ &\Rightarrow e \in A \end{aligned}$$

so we have A=A'. Further

$$\begin{aligned} e \in B &\Rightarrow (e, \{\emptyset\}) \in B \times \{\{\emptyset\}\} \\ &\Rightarrow (e, \{\emptyset\}) \in \langle A, B\rangle \\ &\Rightarrow (e, \{\emptyset\}) \in \langle A', B'\rangle \\ &\Rightarrow (e, \{\emptyset\}) \in B' \times \{\{\emptyset\}\} \text{ as } (e, \{\emptyset\}) \notin A' \times \{\emptyset\} \\ &\Rightarrow e \in B' \\ e \in B' &\Rightarrow (e, \{\emptyset\}) \in B' \times \{\{\emptyset\}\} \\ &\Rightarrow (e, \{\emptyset\}) \in \langle A', B'\rangle \\ &\Rightarrow (e, \{\emptyset\}) \in \langle A, B\rangle \\ &\Rightarrow (e, \{\emptyset\}) \in B \times \{\{\emptyset\}\} \text{ as } (e, \{\emptyset\}) \notin A \times \{\emptyset\} \\ &\Rightarrow e \in B \end{aligned}$$

so we have that $B = B'$

⇐. Assume now that $A = A' \wedge B = B'$ then

$$\begin{aligned} e \in \langle A, B\rangle &\Leftrightarrow e \in (A \times \{\emptyset\}) \bigcup (B \times \{\{\emptyset\}\}) \\ &\Leftrightarrow (e = (f, \emptyset) \wedge f \in A) \vee (e = (g, \{\emptyset\}) \wedge g \in B) \\ &\Leftrightarrow (e = (f, \emptyset) \wedge f \in A') \vee (e = (g, \{\emptyset\}) \wedge g \in B') \\ &\Leftrightarrow e \in (A' \times \{\emptyset\}) \bigcup (B' \times \{\{\emptyset\}\}) \\ &\Leftrightarrow e \in \langle A', B'\rangle \end{aligned}$$

□

Definition 1.82. *A triple* $\langle A, B, C\rangle$ *where* A, B, C *are classes is the class defined by* $(A \times \{\emptyset\}) \bigcup (B \times \{\{\emptyset\}\}) \bigcup (C \times \{\{\{\emptyset\}\}\})$ *(which is a class because* $\emptyset$ *is a set (see 1.76) so* $\{\emptyset\}, \{\{\emptyset\}\}$ *and* $\{\{\{\emptyset\}\}\}$ *are sets (see 1.63) thus* $\langle A, B, C\rangle$ *is a class).*

Theorem 1.83. *If* (A, B, C) *and* (A', B', C') *are triples then* $\langle A, B, C\rangle = \langle A', B', C'\rangle$ *iff* $A = A' \wedge B = B' \wedge C = C'$

Proof.

⇒. Assume $\langle A, B, C\rangle = \langle A', B', C'\rangle$ then we have

$$\begin{aligned}
e \in A &\Rightarrow (e, \emptyset) \in A \times \{\emptyset\} \\
&\Rightarrow (e, \emptyset) \in \langle A, B, C\rangle \\
&\Rightarrow (e, \emptyset) \in \langle A', B', C'\rangle \\
&\Rightarrow (e, \emptyset) \in A' \times \{\emptyset\} \text{ as } (e, \emptyset) \notin B' \times \{\{\emptyset\}\}, C' \times \{\{\{\emptyset\}\}\} \\
&\Rightarrow e \in A' \\
e \in A' &\Rightarrow (e, \emptyset) \in A' \times \{\emptyset\} \\
&\Rightarrow (e, \emptyset) \in \langle A', B', C'\rangle \\
&\Rightarrow (e, \emptyset) \in \langle A, B, C\rangle \\
&\Rightarrow (e, \emptyset) \in A \times \{\emptyset\} \text{ as } (e, \emptyset) \notin B \times \{\{\emptyset\}\}, C \times \{\{\{\emptyset\}\}\} \\
&\Rightarrow e \in A
\end{aligned}$$

So we have $A = A'$

$$\begin{aligned}
e \in B &\Rightarrow (e, \{\emptyset\}) \in B \times \{\{\emptyset\}\} \\
&\Rightarrow (e, \{\emptyset\}) \in \langle A, B, C\rangle \\
&\Rightarrow (e, \{\emptyset\}) \in \langle A', B', C'\rangle \\
&\Rightarrow (e, \{\emptyset\}) \in B' \times \{\{\emptyset\}\} \text{ as } (e, \{\emptyset\}) \notin A' \times \{\emptyset\}, C' \times \{\{\{\emptyset\}\}\} \\
&\Rightarrow e \in B' \\
e \in B' &\Rightarrow (e, \{\emptyset\}) \in B' \times \{\{\emptyset\}\} \\
&\Rightarrow (e, \{\emptyset\}) \in \langle A', B', C'\rangle \\
&\Rightarrow (e, \{\emptyset\}) \in \langle A, B, C\rangle \\
&\Rightarrow (e, \{\emptyset\}) \in B \times \{\{\emptyset\}\} \text{ as } (e, \{\emptyset\}) \notin A \times \{\emptyset\}, C \times \{\{\{\emptyset\}\}\} \\
&\Rightarrow e \in B
\end{aligned}$$

So we have $B = B'$

$$\begin{aligned}
e \in C &\Rightarrow (e, \{\{\emptyset\}\}) \in C \times \{\{\{\emptyset\}\}\} \\
&\Rightarrow (e, \{\{\emptyset\}\}) \in \langle A, B, C\rangle \\
&\Rightarrow (e, \{\{\emptyset\}\}) \in \langle A', B', C'\rangle \\
&\Rightarrow (e, \{\{\emptyset\}\}) \in C' \times \{\{\{\emptyset\}\}\} \text{ as } (e, \{\{\emptyset\}\}) \notin A \times \{\emptyset\}, B \times \{\{\emptyset\}\} \\
&\Rightarrow e \in C' \\
e \in C' &\Rightarrow (e, \{\{\emptyset\}\}) \in C' \times \{\{\{\emptyset\}\}\} \\
&\Rightarrow (e, \{\{\emptyset\}\}) \in \langle A', B', C'\rangle \\
&\Rightarrow (e, \{\{\emptyset\}\}) \in \langle A, B, C\rangle \\
&\Rightarrow (e, \{\{\emptyset\}\}) \in C \times \{\{\{\emptyset\}\}\} \\
&\Rightarrow e \in C
\end{aligned}$$

So we have $C = C'$

⇐. Assume now that $A = A' \wedge B = B' \wedge C = C'$ then

$$\begin{aligned} e \in \langle A, B, C\rangle &\Leftrightarrow e \in (A \times \{\emptyset\}) \bigcup (B \times \{\{\emptyset\}\}) \bigcup (C \times \{\{\{\emptyset\}\}\}) \\ &\Leftrightarrow (e = (f, \emptyset) \wedge f \in A) \vee (e = (g, \{\emptyset\}) \wedge g \in B) \vee (e = (f, \\ &\quad \{\{\emptyset\}\}) \wedge f \in C) \\ &\Leftrightarrow (e = (f, \emptyset) \wedge f \in A') \vee (e = (g, \{\emptyset\}) \wedge g \in B') \vee (e = (f, \\ &\quad \{\{\emptyset\}\}) \wedge f \in C') \\ &\Leftrightarrow e \in (A' \times \{\emptyset\}) \bigcup (B' \times \{\{\emptyset\}\}) \bigcup (C' \times \{\{\{\emptyset\}\}\}) \\ &\Leftrightarrow e \in \langle A', B', C'\rangle \end{aligned}$$

□

1.7 Families

The concept of a family of mathematical objects is central in mathematics. Families are defined in many cases based on functions but for a a family of classes (or later sets) we only need the concept of graphs.

Definition 1.84. *A family of classes is a pair $\langle G, I\rangle$ where G is a graph with* $\mathrm{dom}(G) \subseteq I$*. If $\langle G, I\rangle$ is a family then we define $\forall i \in I$ that $G_i = \{x | (i, x) \in G\}$. As the G_i are the most important objects of a family we note a family as $\{G_i\}_{i \in I}$ instead of $\langle G, I\rangle$. The graph of a family $\{G_i\}_{i \in I}$ is noted as* $\mathrm{graph}(\{G_i\}_{i \in I})$ *and I is the index of the family. So we must have* $\mathrm{dom}(\mathrm{graph}(\{G_i\}_{i \in I})) \subseteq I$*.*

Notice that in this notation a index variable is introduced which is not really needed (which is a disadvantage of this kind of notations) so $\{G_i\}_{i \in I} = \{G_j\}_{j \in I}$.

Example 1.85. If $\{G_i\}_{i \in I}$ is a family with $I = \emptyset$ then $\{G_i\}_{i \in I} = \langle G, I\rangle$ where G is a graph and $\mathrm{dom}(G) \subseteq \emptyset \Rightarrow \mathrm{dom}(G) = \emptyset$. If now $(x, y) \in G$ then $x \in \mathrm{dom}(G) = \emptyset$ a contradiction so we must have then $G = \emptyset$ or hence $\{G_i\}_{i \in I} = \langle \emptyset, \emptyset\rangle$.

Example 1.86. If $G = \{(1, a), (1, b), (1, c), (2, c), (2, d), (3, a)\}$ then we can write G as $\{G_i\}_{i \in I}$ where $I = \{1, 2, 3\}$ and $G_1 = \{a, b, c\}, G_2 = \{c, d\}$ and $G_3 = \{a\}$

In the above example we have that $\mathrm{dom}(\mathrm{graph}(\{G_i\}_{i \in I})) = I$ which for general families is not needed. The reason of this generalization is that we can allow empty G_i's as the following theorem illustrates.

Theorem 1.87. *If $\{G_i\}_{i \in I}$ is a family of classes then we have* $\mathrm{dom}(\mathrm{graph}(\{G_i\}_{i \in I})) = I \Leftrightarrow \forall i \in I \vDash G_i \neq \emptyset$

Proof.

(⇒) If $\mathrm{dom}(\mathrm{graph}(\{G_i\}_{i \in I})) = I$ then if $i \in I$ there exists a x such that $(i, x) \in \mathrm{graph}(\{G_i\}_{i \in I})$ and thus $x \in G_i \Rightarrow G_i \neq \emptyset$.

(⇐) If $\forall i\in I\vDash G_i\neq\emptyset$ then if $i\in I$ there exists a $x\in G_i\Rightarrow(i,x)\in\text{graph}(\{G_i\}_{i\in I})\Rightarrow i\in\text{dom}(\text{graph}(\{G_i\}_{i\in I}))$ proving that $I\subseteq\text{dom}(\text{graph}(\{G_i\}_{i\in I}))$ and as by definition we have $\text{dom}(\text{graph}(\{G_i\}_{i\in I}))\subseteq I$ we have $\text{dom}(\text{graph}(\{G_i\}_{i\in I}))=I$. □

The following lemmas show how to build new families of classes based on a existing family

Lemma 1.88. *Let $\{G_i)_{i\in I}$ be a family of classes with $G=\text{graph}(\{G\}_{i\in I})$ then $\langle F,I\rangle$ where $F=\{(i,x)|i\in I\wedge(i,x)\notin G\}$ is a family with $\forall i\in F_i=G_i^c$. Hence $\langle F,I\rangle$ can be written as $\{G_i^c\}_{i\in I}$.*

Proof. If $i\in\text{dom}(F)$ then there exists a x such that $i\in I\wedge(i,x)\notin G\Rightarrow i\in I$ proving that $\text{dom}(F)\subseteq I$. So $\langle F,I\rangle$ is a family of sets. Take now $i\in I$ then

$$\begin{aligned} x\in F_i &\Leftrightarrow (i,x)\in F\\ &\Leftrightarrow (i,x)\notin G\\ &\Leftrightarrow x\notin G_i\\ &\Leftrightarrow x\in G^c \end{aligned}$$

proving that $F_i=G_i^c$ □

Lemma 1.89. *Let $\{G_i)_{i\in I}$ be a family of classes with $G=\text{graph}(\{G\}_{i\in I})$ and A a class then $\langle F,I\rangle$ where $F=\{(i,x)|(i,x)\in G\wedge x\in A\}\subseteq G$ is a family with $\forall i\in F_i=G_i\bigcap A$. Hence $\langle F,I\rangle$ can be written as $\{G_i\bigcap A\}_{i\in I}$.*

Proof. If $i\in\text{dom}(F)$ then there exists a x such that $(i,x)\in F\subseteq G$ so that $i\in\text{dom}(G)\subseteq I$ proving that $\text{dom}(F)\subseteq I$. So $\langle F,I\rangle$ is a family of sets. Take now $i\in I$ then

$$\begin{aligned} x\in F_i &\Leftrightarrow (i,x)\in F\\ &\Leftrightarrow (i,x)\in G\wedge x\in A\\ &\Leftrightarrow x\in G_i\wedge x\in A\\ &\Leftrightarrow x\in G_i\bigcap A \end{aligned}$$

proving that $F_i=G_i\bigcap A$ □

Lemma 1.90. *Let $\{G_i\}_{i\in I}$ be a family of classes with $G=\text{graph}(\{G\}_{i\in I})$ and A a class then $\langle F,I\rangle$ where $F=G\bigcup(I\times A)$ is a family with $\forall i\in F_i=G_i\bigcup A$. Hence $\langle F,I\rangle$ can be written as $\{G_i\bigcup A\}_{i\in I}$.*

Proof. If $i\in\text{dom}(F)$ then there exists a x such that $(i,x)\in G\bigcup(I\times A)\Rightarrow(i,x)\in G\vee(i,x)\in I\times A\Rightarrow(i\in\text{dom}(G)\subseteq I)\vee i\in I\Rightarrow i\in I$ hence $\text{dom}(F)\subseteq I$. So $\langle F,I\rangle$ is a family of sets. Take now $i\in I$ then

$$\begin{aligned} x\in F_i &\Leftrightarrow (i,x)\in F\\ &\Leftrightarrow (i,x)\in G\vee(i,x)\in I\times A\\ &\underset{i\in I}{\Leftrightarrow} x\in G_i\wedge x\in A\\ &\Leftrightarrow x\in G_i\bigcup A \end{aligned}$$

proving that $F_i = G_i \bigcup A$ □

Lemma 1.91. *Let $\{G_i)_{i \in I}$ be a family of classes with $G = \operatorname{graph}(\{G)_{i \in I})$ and A a class then $\langle F, I \rangle$ where $F = \{(i, x) | (i, x) \in G \wedge x \notin A\} \subseteq G$ is a family with $\forall i \in F_i = G_i \backslash A$. Hence $\langle F, I \rangle$ can be written as $\{G_i \backslash A\}_{i \in I}$.*

Proof. That F is a graph follows from $F \subseteq G \subseteq \mathcal{U} \times \mathcal{U}$ (as G is a graph). If $i \in \operatorname{dom}(F)$ then there exists a x such that $(i, x) \in F \subseteq G$ so that $i \in \operatorname{dom}(G) \subseteq I$ proving that $\operatorname{dom}(F) \subseteq I$. So $\langle F, I \rangle$ is a family of sets. Take now $i \in I$ then

$$\begin{aligned} x \in F_i &\Leftrightarrow (i, x) \in F \\ &\Leftrightarrow (i, x) \in G \wedge x \notin A \\ &\Leftrightarrow x \in G_i \wedge x \notin A \\ &\Leftrightarrow x \in G_i \backslash A \end{aligned}$$

proving that $F_i = G_i \backslash A$ □

Lemma 1.92. *Let $\{G_i\}_{i \in I}$ be a family of classes with $G = \operatorname{graph}(\{G\}_{i \in I})$ and A a class then $\langle F, I \rangle$ where $F = (I \times A) \backslash G$ is a family with $\forall i \in F_i = A \backslash G_i$. Hence $\langle F, I \rangle$ can be written as $\{A \backslash G_i\}_{i \in I}$.*

Proof. If $i \in \operatorname{dom}(F)$ then there exists a x such that $(i, x) \in F \Rightarrow (i, x) \in (I \times A) \backslash G \Rightarrow (i, x) \in I \times A \Rightarrow i \in I$ proving that $\operatorname{dom}(F) \subseteq I$. So $\langle F, I \rangle$ is a family of sets. Take now $i \in I$ then

$$\begin{aligned} x \in F_i &\Leftrightarrow (i, x) \in F \\ &\Leftrightarrow (i, x) \in I \times A \wedge (i, x) \notin G \\ &\underset{i \in I}{\Leftrightarrow} x \in A \wedge x \notin G \\ &\Leftrightarrow x \in A \backslash G_i \end{aligned}$$

proving that $F_i = A \backslash G_i$ □

Lemma 1.93. *Let $\{G_i)i \in I$ be a family of classes with $G = \operatorname{graph}(\{G)_{i \in I})$ and A a class then $\langle F, I \rangle$ where $F = \{(i, (x, a)) | (i, x) \in G \wedge a \in A\}$ is a family with $\forall i \in F_i = G_i \times A$. Hence $\langle F, I \rangle$ can be written as $\{G_i \times A\}_{i \in I}$.*

Proof. If $i \in \operatorname{dom}(F)$ then $\exists z$ such that $z = (i, (x, a))$ and $(i, x) \in G \wedge a \in A \Rightarrow i \in \operatorname{dom}(G) \in I$ proving that $\operatorname{dom}(F) \subseteq I$. So $\langle F, I \rangle$ is a family of sets. Take now $i \in I$ then

$$\begin{aligned} z \in F_i &\Leftrightarrow (i, z) \in F \\ &\Leftrightarrow z = (x, a) \wedge (i, x) \in G \wedge a \in A \\ &\Leftrightarrow z = (x, a) \wedge x \in G_i \wedge a \in A \\ &\Leftrightarrow z \in G_i \times A \end{aligned}$$

proving that $F_i = G_i \times A$ □

Lemma 1.94. *Let $\{G_i\}_{i \in I}$ be a family of classes with $G = \operatorname{graph}(\{G)_{i \in I})$ and A a class then $\langle F, I \rangle$ where $F = \{(i, (a, x)) | (i, x) \in G \wedge a \in A\}$ is a family with $\forall i \in F_i = A \times G_i$. Hence $\langle F, I \rangle$ can be written as $\{A \times G_i\}_{i \in I}$.*

Proof. If $i \in \operatorname{dom}(F)$ then $\exists z$ such that $z = (i, (a, x))$ and $(i, x) \in G \wedge a \in A \Rightarrow i \in \operatorname{dom}(G) \in I$ proving that $\operatorname{dom}(F) \subseteq I$. So $\langle F, I \rangle$ is a family of sets. Take now $i \in I$ then

$$\begin{aligned} z \in F_i &\Leftrightarrow (i, z) \in F \\ &\Leftrightarrow z = (a, x) \wedge (i, x) \in G \wedge a \in A \\ &\Leftrightarrow z = (a, x) \wedge x \in G_i \wedge a \in A \\ &\Leftrightarrow z \in A \times G_i \end{aligned}$$

proving that $F_i = A \times G_i$ □

Lemma 1.95. *Let I, J be classes and $\{A_{(i,j)}\}_{(i,j) \in I \times J}$ a family where A=graph$(\{A_{(i,j)}\}_{(i,j) \in I \times J})$ then we have*

1. *If $i \in I$ then $\langle F, J \rangle$ where $F = \{(j, x) | ((i, j), x) \in A\}$ is a graph and $\forall j \in J$ we have $F_j = A_{(i,j)}$ so that we can write $\langle F, I \rangle$ as $\{A_{(i,j)}\}_{j \in J}$*
2. *If $j \in J$ then $\langle F, I \rangle$ where $F = \{(i, x) | ((i, j), x) \in A\}$ is a graph and $\forall i \in I$ we have $F_i = A_{(i,j)}$ so that we can write $\langle F, I \rangle$ as $\{A_{(i,j)}\}_{i \in I}$*

Proof.

1. If $j \in \operatorname{dom}(F)$ then there exists a x such that $(j, x) \in F$ so that $((i, j), x) \in A$ hence $(i, j) \subseteq I \times J \Rightarrow j \in J$ proving that $\langle F, J \rangle$ is a family. Further if $j \in J$ then

$$\begin{aligned} x \in F_j &\Leftrightarrow (j, x) \in F \\ &\Leftrightarrow ((i, j), x) \in A \\ &\Leftrightarrow x \in A_{(i,j)} \end{aligned}$$

   proving that $F_j = A_{(i,j)}$

2. If $i \in \operatorname{dom}(F)$ then there exists a x such that $(j, x) \in F$ so that $((i, j), x) \in A$ hence $(i, j) \subseteq I \times J \Rightarrow i \in I$ proving that $\langle F, I \rangle$ is a family. Further if $j \in I$ then

$$\begin{aligned} x \in F_j &\Leftrightarrow (j, x) \in F \\ &\Leftrightarrow ((i, j), x) \in A \\ &\Leftrightarrow x \in A_{(i,j)} \end{aligned}$$

   proving that $F_j = A_{(i,j)}$ □

Lemma 1.96. *Let $\{G_i\}_{i \in I}$, $\{F_j\}_{j \in J}$ be families with $G = \operatorname{graph}(\{G_i\}_{i \in I})$ and $F = \operatorname{graph}(\{F_i\}_{i \in I})$ then*

1. *$\langle H, I \times J \rangle$ with $H = \{((i, j), x) | (i, x) \in G \wedge (j, x) \in F\}$ defines a family with $\forall (i, j) \in I \times J$ $H_{(i,j)} = G_i \bigcap F_i$. Hence $(H, I \times J)$ can be written as $\{H_i \bigcap F_j\}_{(i,j) \in I \times J}$*
2. *$\langle H, I \times J \rangle$ with $H = \{((i, j), x) | (i, j) \in I \times J \wedge ((i, x) \in G \vee (i, x) \in F)\}$ defines a family with $\forall (i, j) \in I \times J$ $H_{(i,j)} = G_i \bigcup F_i$. Hence $(H, I \times J)$ can be written as $\{H_i \bigcup F_j\}_{(i,j) \in I \times J}$*

Proof.

1. If $z \in \operatorname{dom}(H)$ then $z = (i, j)$ and there exists a x such that $(i, x) \in G \wedge (j, x) \in F \Rightarrow i \in \operatorname{dom}(G) \subseteq I \wedge j \in \operatorname{dom}(F) \subseteq J$ hence $z = (i, j) \in I \times J$ proving that $\operatorname{dom}(H) \subseteq I \times J$. So $\langle H, I \times J\rangle$ is a family. Further if $z = (i, j) \in \operatorname{dom}(H)$ then we have

$$\begin{aligned} x \in H_z &\Leftrightarrow ((i, j, x) \in H \\ &\Leftrightarrow (i, x) \in G \wedge (j, x) \in F \\ &\Leftrightarrow x \in G_i \wedge y \in F_j \\ &\Leftrightarrow x \in G_i \bigcap F_j \end{aligned}$$

   proving that $H_z = H_{(i,j)} = G_i \bigcap H_j$

2. If $z \in \operatorname{dom}(H)$ then $z = (i, j) \in I \times J \wedge ((i, x) \in G \vee (j, y) \in F) \Rightarrow (i, j) \in I \times J$ proving $\operatorname{dom}(H) \subseteq I \times J$. So $\langle H, I\rangle$ is a family. If $z = (i, j) \in I \times J$ then

$$\begin{aligned} x \in H_z &\Leftrightarrow ((i, j), x) \in H \\ &\Leftrightarrow (i, x) \in G \vee (j, x) \in F_j \\ &\Leftrightarrow x \in G_i \vee x \in F_j \\ &\Leftrightarrow x \in G_i \bigcup F_j \end{aligned}$$

   proving that proving that $H_z = H_{(i,j)} = G_i \bigcap H_j$

$\square$

To summarize the above lemmas we have given families $\{G_i\}_{i \in I}$, $\{F_j\}_{j \in I}$ and a class A constructed the following new families:

1. $\{G_i^c\}_{i \in I}$
2. $\{G_i \bigcap A\}_{i \in I}$
3. $\{G_i \bigcup A\}_{i \in I}$
4. $\{G_i \setminus A\}_{i \in I}$
5. $\{A \setminus G\}_{i \in I}$
6. $\{G_i \bigcap F_j\}_{(i,j) \in I \times J}$
7. $\{G_i \bigcup F_j\}_{(i,j) \in I \times J}$

One of the most used operations on a family of classes is calculating the union and intersection of the classes in the family.

Definition 1.97. *If* $\{G_i\}_{i \in I}$ *is a indexed family of classes then* $\bigcup_{i \in I} G_i = \{x | \exists i \in I \vdash x \in G_i\}$ *or written using the graph and index we have* $\bigcup_{i \in I} G_i = \{x | \exists i \vdash i \in I \wedge (i, x) \in \operatorname{graph}(\{G_i\}_{i \in I})\}$

Definition 1.98. *If* $\{G_i\}_{i \in I}$ *is a indexed family of classes then* $\bigcap_{i \in I} G_i = \{x | \forall i \in I \vDash x \in G_i\}$ *or written using the graph and index we have* $\bigcap_{i \in I} G_i = \{x | \forall i \in I \vDash (i, x) \in \operatorname{graph}(\{G_i\}_{i \in I})\}$

The following example proves that the normal definition of union and intersection corresponds with the definition of a union and intersection of graphs.

Example 1.99. If A and B are classes define $I = \{0, 1\}$ where $0, 1$ are different elements (we will see later that they exists) and $G = (\{0\} \times A) \bigcup (\{1\} \times B)$ we have then that

1. $G = \{G_i\}_{i \in \{0,1\}}$ is a indexed family of classes with $G_0 = A$ and $G_1 = B$
2. $\bigcup_{i \in \{0,1\}} G_i = A \bigcup B$
3. $\bigcap_{i \in \{0,1\}} G_i = A \bigcap B$

Proof. First we must prove that $\operatorname{dom}(G) \subseteq \{0, 1\}$. If $x \in \operatorname{dom}(G) \Rightarrow \exists y \vdash (x, y) \in G \Rightarrow (x, y) \in \{0\} \times A \vee (x, y) \in \{1\} \times B \Rightarrow x \in \{0\} \vee x \in \{1\} \Rightarrow x = 0 \vee x = 1 \Rightarrow x \in \{0, 1\} \Rightarrow \operatorname{dom}(G) \subseteq \{0, 1\}$. So we have that G is a family indexed by $I = \{0, 1\}$. Now we have

1.

$$\begin{aligned} x \in G_0 &\Leftrightarrow (0, x) \in (\{0\} \times A) \bigcup (\{1\} \times B) \\ &\Leftrightarrow (0, x) \in \{0\} \times A \vee (0, x) \in \{1\} \times B \\ &\Leftrightarrow (0, x) \in \{0\} \times A \vee (0 = 1 \wedge x \in B) \\ &\underset{0 \neq 1}{\Leftrightarrow} (0, x) \in \{0\} \times A \\ &\Leftrightarrow x \in A \end{aligned}$$

and

$$\begin{aligned} x \in G_1 &\Leftrightarrow (1, x) \in (\{0\} \times A) \bigcup (\{1\} \times B) \\ &\Leftrightarrow (1, x) \in \{0\} \times A \vee (1, x) \in \{1\} \times B \\ &\Leftrightarrow (1 = 0 \wedge x \in A) \vee (1, x) \in \{1\} \times B \\ &\underset{0 \neq 1}{\Leftrightarrow} (1, x) \in \{1\} \times B \\ &\Leftrightarrow x \in B \end{aligned}$$

2.

$$\begin{aligned} x \in \bigcap_{i \in \{0,1\}} G_i &\Leftrightarrow \forall i \in I \vDash x \in G_i \\ &\Leftrightarrow x \in G_0 \wedge x \in G_1 \\ &\Leftrightarrow x \in A \wedge x \in B \\ &\Leftrightarrow x \in A \bigcap B \end{aligned}$$

3.

$$\begin{aligned} x \in \bigcup_{i \in \{0,1\}} G_i &\Leftrightarrow \exists i \in I \vdash x \in G_i \\ &\Leftrightarrow x \in G_0 \vee x \in G_1 \\ &\Leftrightarrow x \in A \vee x \in B \\ &\Leftrightarrow x \in A \bigcup B \end{aligned}$$

□

Example 1.100. If $\{G_i\}_{i \in \emptyset} = \langle \emptyset, \emptyset \rangle$ (see 1.85) then we have

1. $\bigcup_{i \in \emptyset} G_i = \emptyset$
2. $\bigcap_{i \in \emptyset} G_i = \mathcal{U}$

Proof.

1. If $x \in \bigcup_{i \in \emptyset} G_i$ then $\exists i \in I$ such that $x \in G_i$ which as $I = \emptyset$ is impossible and thus $\bigcup_{i \in \emptyset} G_i = \emptyset$
2. $\bigcap_{i \in \emptyset} G_i = \{x | \forall i \in I \vDash x \in G_i\} \underset{\text{defined}}{=} \{x | x \text{ is } a \text{ element} \wedge \forall i \in I \vDash x \in G_i\}$. As $\bigcap_{i \in \emptyset} G_i \subseteq \mathcal{U}$ we must prove that $\mathcal{U} \subseteq \bigcap_{i \in \emptyset} G_i$ so if $x \in \mathcal{U}$ then x is a element $\Rightarrow x$ is a element $\wedge$ $(\forall i \in \emptyset \vDash x \in G_i)$ [is vacuously true] so that $x \in \bigcap_{i \in \emptyset} G_i$. $\square$

As the only defining elements for classes are classes and the relation 'is a element of' we can consider a class itself as a collection of classes (these classes are then elements and thus by definition sets). This gives rise to another type of general union and intersection.

Definition 1.101. *If $\mathcal{A}$ is a class then $\bigcap_{A \in \mathcal{A}} A = \{x | \forall A \in \mathcal{A} \vDash x \in A\}$ and $\bigcup_{A \in \mathcal{A}} = \{x | \exists A \in \mathcal{A} \vDash x \in A\}$*

We examine now the relation of the above definition and the union (intersection) of families. Let $\mathcal{A}$ be a class of classes and define $\mathcal{I}^{\mathcal{A}} = \{(A, x) | A \in \mathcal{A} \wedge x \in A\}$ then $\mathcal{I}^{\mathcal{A}} \subseteq \mathcal{A} \times \bigcup_{A \in \mathcal{A}} A \subseteq \mathcal{U} \times \mathcal{U}$ so that $\mathcal{I}^{\mathcal{A}}$ is a graph with $\operatorname{dom}(\mathcal{I}^{\mathcal{A}}) \subseteq \mathcal{A}$. Also if $A \in \mathcal{A}$ then $(\mathcal{I}^{\mathcal{A}})_A = \{x | (A, x) \in \mathcal{I}^{\mathcal{A}}\} = \{x | A \in \mathcal{A} \wedge x \in A\} \underset{A \in \mathcal{A} \text{is true}}{=} \{x | x \in A\} = A$ so that the family $\{(\mathcal{I}^{\mathcal{A}})_A\}_{A \in \mathcal{A}}$ can be be written as $\{A\}_{A \in \mathcal{A}}$. Essentially we can consider a class $\mathcal{A}$ as a family $\{(\mathcal{I}^{\mathcal{A}})_A\}_{A \in \mathcal{A}}$ with $A \in \mathcal{A} \Leftrightarrow (\mathcal{I}^{\mathcal{A}})_A = A$, that is we index the class by itself. This leads to the following definition:

Definition 1.102. *If $\mathcal{A}$ is a class of classes then $\{A\}_{A \in \mathcal{A}}$ is equal to $\{A_i\}_{i \in \mathcal{A}}$ where* $\operatorname{graph}(\{A_i\}_{i \in \mathcal{A}}) = \mathcal{I}^{\mathcal{A}} = \{(A, x) | A \in \mathcal{A} \wedge x \in A\}$

The following theorem shows that the two definitions of unions and intersection are essentially the same.

Theorem 1.103. *If $\mathcal{A}$ is a class then $\bigcup_{A \in \mathcal{A}} (\mathcal{I}^{\mathcal{A}})_A = \bigcup_{A \in \mathcal{A}} A$ and $\bigcap_{A \in \mathcal{A}} (\mathcal{I}^{\mathcal{A}})_A = \bigcap_{A \in \mathcal{A}} A$*

Proof.

$$
\begin{aligned}
x \in \bigcup_{A \in \mathcal{A}} (\mathcal{I}^{\mathcal{A}})_A &\Leftrightarrow \exists A \in \mathcal{A} \vdash x \in (\mathcal{I}^{\mathcal{A}})_A \\
&\underset{(\mathcal{I}^{\mathcal{A}})_A = A}{\Leftrightarrow} \exists A \in \mathcal{A} \vdash x \in A \\
&\Leftrightarrow x \in \bigcup_{A \in A} A \\
x \in \bigcap_{A \in \mathcal{A}} (\mathcal{I}^{\mathcal{A}})_A &\Leftrightarrow \forall A \in \mathcal{A} \vdash x \in (\mathcal{I}^{\mathcal{A}})_A \\
&\underset{(\mathcal{I}^{\mathcal{A}})_A = A}{\Leftrightarrow} \forall A \in \mathcal{A} \vdash x \in A \\
&\Leftrightarrow x \in \bigcap_{A \in \mathcal{A}} A
\end{aligned}
$$

$\square$

Theorem 1.104. *If $\{G_i\}_{i\in I}$ is a family of classes and B is a class then we have*

1. $\forall i\in I\vdash G_i\subseteq\bigcup_{i\in I}G_i$
2. $\forall i\in I\vdash\bigcap_{i\in I}G_i\subseteq G_i$
3. *If* $\forall i\in I$ *we have* $A_i\subseteq B$ *then* $\bigcup_{i\in I}A_i\subseteq B$
4. *If* $\forall i\in I$ *we have* $B\subseteq A_i$ *then* $B\subseteq\bigcap A_{i\in I}$

Proof.

1. *Take* $i\in I$ *then if* $x\in G_i\Rightarrow\exists j\vdash x\in G_j\,(\text{take}\,i=j)\Rightarrow x\in\bigcup_{i\in I}G_i$
2. *Take* $i\in I$ *then if* $x\in\bigcap_{j\in I}G_j\Rightarrow\forall j\in I\vdash x\in G_j\Rightarrow x\in G_i$
3. *We have*

$$\begin{aligned} x\in\bigcup_{i\in I}A_i &\Rightarrow \exists i\in I\vdash x\in A_i\\ &\underset{A_i\subseteq B}{\Rightarrow} x\in B\\ &\Rightarrow \bigcup_{i\in I}A_i\subseteq B \end{aligned}$$

4. *We have*

$$\begin{aligned} x\in B &\Rightarrow \forall i\in I\vDash x\in A_i\\ &\Rightarrow x\in\bigcap_{i\in I}A_i\\ &\Rightarrow B\subseteq\bigcap_{i\in I}A_i \end{aligned}$$

□

Theorem 1.105. (Generalized deMorgan's Laws) *Let $\{G_i\}_{i\in I}$ be an indexed family of classes. Then*

1. $(\bigcup_{i\in I}A_i)^c=\bigcap_{i\in I}A_i^c$
2. $(\bigcap_{i\in I}A_i)^c=\bigcup_{i\in I}A_i^c$

Proof.

1. This is proved by

$$\begin{aligned} x\in\left(\bigcup_{i\in I}A_i\right)^c &\Leftrightarrow x\notin\bigcup_{i\in I}A_i\\ &\Leftrightarrow \neg(\exists i\in I\vdash x\in A_i)\\ &\Leftrightarrow \forall i\in I\vDash x\notin A_i\\ &\Leftrightarrow \forall i\in I\vDash x\in A_i^c\\ &\Leftrightarrow x\in\bigcap_{i\in I}A_i^c \end{aligned}$$

2. This is proved by

$$\begin{aligned} x \in \left(\bigcap_{i\in I} A_i\right)^c &\Leftrightarrow x \notin \bigcap_{i\in I} A_i \\ &\Leftrightarrow \neg(\forall i \in I \vDash x \in A_i) \\ &\Leftrightarrow \exists i \in I \vDash x \notin A_i \\ &\Leftrightarrow \exists i \in I \vDash x \in A_i^c \\ &\Leftrightarrow x \in \bigcup_{i\in I} A_i^c \end{aligned}$$

□

Theorem 1.106. *Let $\{A_{(i,j)}\}_{(i,j)\in I\times J}$ be a family then*

1. $\bigcup_{(i,j)\in I\times J} A_{(i,j)} = \bigcup_{i\in I} (\bigcup_{j\in J} A_{(i,j)})$
2. $\bigcap_{(i,j)\in I\times J} A_{(i,j)} = \bigcap_{i\in I} (\bigcap_{j\in J} A_{(i,j)})$

Proof.

1.

$$\begin{aligned} x \in \bigcup_{(i,j)\in I\times J} A_{(i,j)} &\Leftrightarrow \exists (i,j) \in I \times J \text{ with } x \in A_{(i,j)} \\ &\Leftrightarrow \exists i \in I \text{ with } \exists j \in J \text{ such that } x \in A_{(i,j)} \\ &\Leftrightarrow \exists i \in I \text{ with } x \in \bigcup_{j\in J} A_{(i,j)} \\ &\Leftrightarrow x \in \bigcup_{i\in I} \left(\bigcup_{j\in J} A_{(i,j)}\right) \end{aligned}$$

2.

$$\begin{aligned} x \in \bigcap_{(i,j)\in I\times J} A_{(i,j)} &\Leftrightarrow \forall (i,j) \in I \times J \,\text{so that}\, x \in A_{(i,j)} \\ &\Leftrightarrow \forall i \in I \,\text{we have}\, \forall j \in J \,\text{we have}\, x \in A_{(i,j)} \\ &\Leftrightarrow \forall i \in I \,\text{we have}\, x \in \bigcap_{j\in J} A_{(i,j)} \\ &\Leftrightarrow x \in \bigcap_{i\in I} \left(\bigcap_{j\in J} A_{(i,j)}\right) \end{aligned}$$

□

Theorem 1.107. *If $\{A_i\}_{i\in I}$ and $\{B_j\}_{j\in J}$ are indexed families of classes and a class C then*

1. $(\bigcup_{i\in I} A_i) \bigcap C = \bigcup_{i\in I} (A_i \bigcap C)$
2. $(\bigcap_{i\in I} A_i) \bigcup C = \bigcap_{i\in I} (A_i \bigcup C)$
3. $(\bigcup_{i\in I} A_i) \bigcap (\bigcup_{j\in J} B_j) = \bigcup_{(i,j)\in I\times J} (A_i \bigcap B_j)$
4. $(\bigcap_{i\in I} A_i) \bigcup (\bigcap_{j\in J} B_j) = \bigcap_{(i,j)\in I\times J} (A_i \bigcup B_j)$
5. $\bigcup_{i\in I} A_i = \left(\bigcup_{i\in I\setminus\{j\}} A_i\right) \bigcup A_j$ *if* $j \in I$

6. $\bigcap_{i\in I} A_i = \left(\bigcap_{i\in I\setminus\{j\}} A_j\right)\bigcap A_j$ *if* $j\in I$
7. $(\bigcup_{i\in I} A_i)\bigcup C = \bigcup_{i\in I}(A_i\bigcup C)$
8. $(\bigcap_{i\in I} A_i)\bigcap C = \bigcap_{i\in I}(A_i\bigcap C)$

Proof.

1. This is proved by

$$\begin{aligned} x\in\left(\bigcup_{i\in I} A_i\right)\bigcap C &\Leftrightarrow x\in\bigcup_{i\in I} A_i\wedge x\in C\\ &\Leftrightarrow (\exists i\in I\vDash x\in A_i)\wedge x\in C\\ &\Leftrightarrow \exists i\in I\vDash x\in A_i\wedge x\in C\\ &\Leftrightarrow \exists i\in I|x\in A_i\bigcap C\\ &\Leftrightarrow x\in\bigcup_{i\in I}\left(A_i\bigcap C\right)\end{aligned}$$

2. This is proved by

$$\begin{aligned} x\in\left(\bigcap_{i\in I} A_i\right)\bigcup C &\Leftrightarrow x\in\bigcap_{i\in I} A_i\vee x\in C\\ &\Leftrightarrow (\forall i\in I\vDash x\in A_i)\vee x\in C\\ &\Leftrightarrow \forall i\in I\vDash x\in A_i\vee x\in C\\ &\Leftrightarrow \forall i\in I\vDash x\in A_i\bigcap C\\ &\Leftrightarrow x\in\bigcap_{i\in I}\left(A_i\bigcap\right)\end{aligned}$$

3. This is proved by

$$\begin{aligned} x\in\left(\bigcup_{i\in I} A_i\right)\bigcap\left(\bigcup_{j\in J} B_j\right) &\Leftrightarrow x\in\bigcup_{i\in I} A_i\wedge x\in\bigcup_{j\in J} B_j\\ &\Leftrightarrow (\exists i\in I\vdash x\in A_i)\wedge(\exists j\in J\vdash x\in B_j)\\ &\Leftrightarrow \exists(i,j)\in I\times J\vdash x\in A_i\wedge x\in B_j\\ &\Leftrightarrow \exists(i,j)\in I\times J\vdash x\in A_i\bigcap B_j\\ &\Leftrightarrow x\in\bigcup_{(i.j)\in I\times J}\left(A_i\bigcap B_j\right)\end{aligned}$$

4. This is proved by

$$\begin{aligned} x\in\left(\bigcap_{i\in I} A_i\right)\bigcup\left(\bigcap_{j\in J} B_j\right) &\Leftrightarrow x\in\bigcap_{i\in I} A_i\vee x\in\bigcap_{j\in J} B_j\\ &\Leftrightarrow (\forall i\in I\vDash x\in A_i)\vee(\forall j\in J\vDash x\in B_j)\\ &\Leftrightarrow \forall(i,j)\in I\times J\vDash(x\in A_i\vee x\in B_j)\\ &\Leftrightarrow \forall(i,j)\in I\times J\vDash x\in A_i\bigcup B_j\\ &\Leftrightarrow x\in\bigcap_{(i,j)\in I\times J}\left(A_i\bigcup B_j\right)\end{aligned}$$

5. If $j \in I \Rightarrow I = (I \backslash \{j\}) \bigcup \{j\}$

$$\begin{aligned} x \in \bigcup_{i \in I} A_i &\Leftrightarrow \exists i \in I \vdash x \in A_i \\ &\Leftrightarrow \exists i \in \big(I \backslash \{j\} \bigcup \{j\} \big) \vdash x \in A_i \\ &\Leftrightarrow (\exists i \in I \backslash \{j\} \vdash x \in A_i) \vee (\exists i \in \{j\} \vdash x \in A_i) \\ &\Leftrightarrow (\exists i \in I \backslash \{j\} \vdash x \in A_i) \vee x \in A_j) \\ &\Leftrightarrow x \in \left(\bigcup_{i \in I \backslash \{j\}} A_i \right) \bigcup A_j \end{aligned}$$

6. If $j \in I \Rightarrow I = (I \backslash \{j\}) \bigcup \{j\}$

$$\begin{aligned} x \in \bigcup_{i \in I} A_i &\Leftrightarrow \forall i \in I \vDash x \in A_i \\ &\Leftrightarrow \forall i \in (I \backslash \{j\}) \bigcup \{j\} \vDash x \in A_i \\ &\Leftrightarrow (\forall i \in (I \backslash \{j\}) \vDash x \in A_i) \wedge x \in A_j \\ &\Leftrightarrow \left(x \in \bigcap_{i \in I \backslash \{j\}} A_j \right) \wedge x \in A_j \\ &\Leftrightarrow x \in \left(\bigcap_{i \in I \backslash \{j\}} A_i \right) \bigcap A_j \end{aligned}$$

7.

$$\begin{aligned} x \in \left(\bigcup_{i \in I} A_i \right) \bigcup C &\Leftrightarrow x \in \bigcup_{i \in I} A_i \vee x \in C \\ &\Leftrightarrow x \in C \vee x \in \bigcup_{i \in} A_i \\ &\Leftrightarrow x \in C \vee (\exists i \in I | x \in A_i) \\ &\Leftrightarrow \exists i \in I \vDash (x \in C \vee x \in A_i) \\ &\Leftrightarrow x \in \bigcup_{i \in I} \big(A_i \bigcup C \big) \end{aligned}$$

8.

$$\begin{aligned} x \in \left(\bigcap_{i \in I} A_i \right) \bigcap C &\Leftrightarrow x \in C \wedge x \in \bigcap_{i \in I} A_i \\ &\Leftrightarrow x \in C \wedge \forall i \in I \vDash x \in A_i \\ &\Leftrightarrow \forall i \in I \vDash (x \in A_i \wedge x \in C) \\ &\Leftrightarrow x \in \left(\bigcap_{i} \big(A_i \bigcap C \big) \right) \end{aligned}$$

$\square$

Theorem 1.108. *If A is a class and $\{A_i\}_{i\in I}$ a family of classes then*

1. $A\setminus(\bigcap_{i\in I}A_i)=\bigcup_{i\in I}(A\setminus A_i)$
2. $A\setminus(\bigcup_{i\in I}A_i)=\bigcap_{i\in I}(A\setminus A_i)$
3. $(\bigcup_{i\in I}A_i)\setminus A=\bigcup_{i\in I}(A_i\setminus A)$
4. $(\bigcap_{i\in I}A_i)\setminus A=\bigcap_{i\in I}(A_i\setminus A)$
5. $A\times(\bigcup_{i\in I}A_i)=\bigcup_{i\in I}(A\times A_i)$
6. $(\bigcup_{i\in I}A_i)\times A=\bigcup_{i\in I}(A_i\times A)$
7. $\bigcap_{i\in I}(A\times A_i)=A\times(\bigcap_{i\in I}A_i)$
8. $\bigcap_{i\in I}(A_i\times A)=(\bigcap_{i\in I}A_i)\times A$

Proof.

1. We have

$$\begin{aligned}
A\setminus\left(\bigcap_{i\in I}A_i\right) &= A\bigcap\left(\bigcap_{i\in I}A_i\right)^c\\
&\underset{1.105}{=} A\bigcap\left(\bigcup_{i\in I}A_i^c\right)\\
&\underset{1.107}{=} \bigcup_{i\in I}\left(A\bigcap A_i^c\right)\\
&= \bigcup_{i\in I}(A\setminus A_i)
\end{aligned}$$

2. We have

$$\begin{aligned}
A\setminus\left(\bigcup_{i\in I}A_i\right) &= A\bigcap\left(\bigcup_{i\in I}A_i\right)^c\\
&= A\bigcap\left(\bigcap_{i\in I}A_i^c\right)\\
&\underset{1.107}{=} \bigcap_{i\in I}\left(A\bigcap A_i^c\right)\\
&= \bigcap_{i\in I}(A\setminus A_i)
\end{aligned}$$

3. $x\in(\bigcup_{i\in I}A_i)\setminus A\Leftrightarrow x\notin A\wedge\exists i\in I\vdash x\in A_i\Leftrightarrow\exists i\in I\vdash(x\notin A\wedge x\in A_i)\Leftrightarrow x\in\bigcup_{i\in I}(A_i\setminus A)$

4. $x \in (\bigcap_{i \in I} A_i) \setminus A \Leftrightarrow x \notin A \wedge \forall i \in I \vdash x \in A_i \Leftrightarrow \forall i \in I \vDash (x \in A_i \wedge x \notin A) \Leftrightarrow x \in \bigcap_{i \in I} (A_i \setminus A)$

5. We have

$$\begin{aligned} (x,y) \in A \times \left(\bigcup_{i \in I} A_i \right) &\Leftrightarrow x \in A \wedge y \in \bigcup_{i \in I} A_i \\ &\Leftrightarrow x \in A \wedge \exists i \in I \vDash y \in A_i \\ &\Leftrightarrow \exists i \in I \vDash x \in A \wedge y \in A_i \\ &\Leftrightarrow (x,y) \in \bigcup_{i \in I} (A \times A_i) \end{aligned}$$

6. We have

$$\begin{aligned} (x,y) \in \left(\bigcup_{i \in I} A_i \right) \times A &\Leftrightarrow x \in \bigcup_{i \in I} A_i \wedge y \in A \\ &\Leftrightarrow y \in A \wedge \exists i \in I \vDash x \in A_i \\ &\Leftrightarrow \exists i \in I \vDash x \in A_i \wedge y \in A \\ &\Leftrightarrow (x,y) \in \bigcup_{i \in I} (A_i \times A) \end{aligned}$$

7. We have

$$\begin{aligned} (x,y) \in \bigcap_{i \in I} (A \times A_i) &\Leftrightarrow \forall i \in I \vDash (x,y) \in A \times A_i \\ &\Leftrightarrow \forall i \in I \vDash x \in A \wedge y \in A_i \\ &\Leftrightarrow x \in A \wedge \forall i \in I \vDash y \in A_i \\ &\Leftrightarrow x \in A \wedge y \in \bigcap_{i \in I} A_i \\ &\Leftrightarrow (x,y) \in A \times \left(\bigcap_{i \in I} A_i \right) \end{aligned}$$

8. We have

$$\begin{aligned} (x,y) \in \bigcap_{i \in I} (A_i \times A) &\Leftrightarrow \forall i \in I \vDash (x,y) \in A_i \times A \\ &\Leftrightarrow \forall i \in I \vDash x \in A_i \wedge y \in A \\ &\Leftrightarrow (\forall i \in I \vDash x \in A_i) \wedge y \in A \\ &\Leftrightarrow x \in \bigcap_{i \in I} A_i \wedge y \in A \\ &\Leftrightarrow (x,y) \in \left(\bigcap_{i \in I} A_i \right) \times A \end{aligned}$$

$\square$

Theorem 1.109. *Let* $\{A_i\}_{i \in I}$ *be a family of classes then* $\bigcup_{i \in I} A_i = \bigcup_{i \in \{j \in I \mid A_j \neq \emptyset\}} A_i$

Proof.

$$\begin{array}{rcl}
x\in\bigcup_{i\in I}A_i & \Rightarrow & \exists i\in I\vdash x\in A_i\\
& \underset{x\in A_i\Rightarrow A_i\neq\emptyset}{\Rightarrow} & \exists i\in\{j\in I\,|\,A_j\neq\emptyset\}\vdash x\in A_i\\
& \Rightarrow & x\in\bigcup_{i\in\{j\in I\,|\,A_j\neq\emptyset\}}A_i\\
x\in\bigcup_{i\in\{j\in I\,|\,A_j\neq\emptyset\}}A_i & \Rightarrow & \exists i\in\{j\in I\,|\,A_j\neq\emptyset\}\vdash x\in A_i\\
& \underset{\{j\in I\,|\,A_j\neq\emptyset\}\subseteq I}{\Rightarrow} & \exists i\in I\vdash x\in A_i\\
& \Rightarrow & x\in\bigcup_{i\in I}A_i
\end{array}$$

$\square$

Theorem 1.110. *Let $\{G_i\}_{i\in I}$ be a family of graphs then*

1. $\mathrm{dom}(\bigcup_{i\in I}G_i)=\bigcup_{i\in I}(\mathrm{dom}(G_i))$

2. $\mathrm{range}(\bigcup_{i\in I}G_i)=\bigcup_{i\in I}(\mathrm{range}(G_i))$

Proof.

1. We have

$$\begin{array}{rcl}
x\in\mathrm{dom}\left(\bigcup_{i\in I}G_i\right) & \Leftrightarrow & \exists y\vdash(x,y)\in\bigcup_{i\in I}G_i\\
& \Leftrightarrow & \exists y\vdash(\exists i\in I\vdash(x,y)\in G_i)\\
& \Leftrightarrow & \exists i\in I\vdash(\exists y\vdash(x,y)\in G_i)\\
& \Leftrightarrow & \exists i\in I\vdash x\in\mathrm{dom}(G_i)\\
& \Leftrightarrow & x\in\bigcup_{i\in I}\mathrm{dom}(G_i)
\end{array}$$

2. This is proved by

$$\begin{array}{rcl}
x\in\mathrm{range}\left(\bigcup_{i\in I}G_i\right) & \Leftrightarrow & \exists y\vdash(y,x)\in\bigcup_{i\in I}G_i\\
& \Leftrightarrow & \exists y\vdash(\exists i\in I\vdash(y,x)\in G_i)\\
& \Leftrightarrow & \exists i\in I\vdash(\exists y\vdash(y,x)\in G_i)\\
& \Leftrightarrow & \exists i\in I\vdash(x\in\mathrm{range}(G_i))\\
& \Leftrightarrow & \bigcup_{i\in I}\mathrm{range}(G_i)
\end{array}$$

$\square$

Chapter 2

Partial functions, functions and relations

2.1 Functions

Functions are the mathematical objects that maps elements of one class to elements of a second class. A function is a triple consisting of a graph representing the mapping, the originating class and the destination class. To make sure that the mapping is deterministic we require that the graph is a function graph so that each element is mapped on only one element. First we define partial functions where not every element of the originating class is mapped and then extend this definition to functions (or mappings as it is sometime called)

2.1.1 Partial functions

Definition 2.1. *A partial function from A to B is a triple of objects $\langle f, A, B\rangle$ where A, B are classes and $f \subseteq A \times B$ [which as $A \times B \subseteq \mathcal{U} \times \mathcal{U}$ means that f is a graph] such that $\forall (x, y), (x, y') \in f$ we have $y = y'$ [or f is a function graph (see 1.60)].*

Notation 2.2. *Instead of writing $\langle f, A, B\rangle$ for a partial function we write $f: A \rightarrow B$*

The following theorem give the necessary conditions needed to ensure that a triple $\langle f, A, B\rangle$ is a partial function

Theorem 2.3. *If f is a graph and A, B are classes then $\langle f, A, B\rangle$ (or $f: A \rightarrow B$) is a partial function iff*

1. $\text{range}(f) \subseteq B$
2. $\text{dom}(f) \subseteq A$
3. *If* $(x, y) \in f \wedge (x, y') \in f \Rightarrow y = y'$ *(or f is a function graph)*

Proof.

$\Rightarrow$. If $\langle f, A, B\rangle$ is a partial function then

1. $y \in \text{range}(f) \Rightarrow \exists x \vdash (x, y) \in f \subseteq A \times B \Rightarrow y \in B \Rightarrow \text{range}(f) \subseteq B$

2. $x \in \mathrm{dom}(f) \Rightarrow \exists y \vdash (x, y) \in f \subseteq A \times B \Rightarrow x \in A \Rightarrow \mathrm{dom}(f) \subseteq A$
3. If $(x, y) \in f \wedge (x, y') \in f \Rightarrow y = y'$ (by definition)

$\Leftarrow$. If $\langle f, A, B\rangle$ fulfills (1),(2) and (3) then $(x, y) \in f \Rightarrow x \in \mathrm{dom}(f) \wedge y \in \mathrm{range}(f) \Rightarrow x \in A \wedge y \in B \Rightarrow (x, y) \in A \times B \Rightarrow f \subseteq A \times B$ and because of (3) we have that f is a function graph. So $\langle f, A, B\rangle$ is a partial function. □

Notation 2.4. *If $f: A \rightarrow B$ is a partial function then for every $x \in \mathrm{dom}(f)$ we have a unique $y \in B$ such that $(x, y) \in f$ we call this y $f(x)$. So $y = f(x)$ is equivalent with $y \in B$ such that $(x, y) \in f$*

Definition 2.5. *Let $f: A \rightarrow B$ be a partial function then if $C \subseteq A$ and $D \subseteq B$ we define $f(C) = \{y \in B | \exists x \in C \vdash (x, y) \in f\}$ (**the image of C by f**) and $f^{-1}(D) = \{x \in A | \exists y \in D \vdash (x, y) \in F\}$ (**the reverse image of D by f**)*

Theorem 2.6. *If $f: A \rightarrow B$ is a partial function and $g: C \rightarrow D$ is a partial function then $g \circ f: A \rightarrow D$ is a partial function.*

Proof. We use the previous theorem (see 2.3) to prove this.

1. Using 1.57 we have that $\mathrm{range}(g \circ f) \subseteq \mathrm{range}(g) \subseteq D$
2. Again using 1.57 we have that $\mathrm{dom}(g \circ f) \subseteq \mathrm{dom}(f) \subseteq A$
3. As g, f are function graphs we have by 1.61 that $g \circ f$ is a function graph. □

Note 2.7. If $f: A \rightarrow B$ and $g: C \rightarrow D$ are partial functions then

$$\begin{aligned} y = (g \circ f)(x) &\underset{\text{notation}}{\Leftrightarrow} (x, y) \in g \circ f \\ &\Leftrightarrow \exists z \in C \text{ such that } (x, z) \in f \wedge (z, y) \in g \\ &\underset{\text{notation}}{\Leftrightarrow} z = f(x) \wedge y = g(z) \\ &\Leftrightarrow y = f(g(z)) \end{aligned}$$

Theorem 2.8. *If $f: A \rightarrow B$ and $g: C \rightarrow D$ are partial functions then $\mathrm{dom}(g \circ f) = \mathrm{dom}(f) \bigcap f^{-1}(\mathrm{dom}(g))$ and $\mathrm{range}(g \circ f) = g(\mathrm{range}(f) \bigcap \mathrm{dom}(g))$ for $g \circ f: A \rightarrow D$*

Proof. If $x \in \mathrm{dom}(g \circ f)$ then there exists a y such that $(x, y) \in g \circ f$, but then there exists a z such that $(x, z) \in f$ and $(z, y) \in g$. From $(x, z) \in f$ it follows that $x \in \mathrm{dom}(f) \subseteq A$ and from (z, y) that $z \in \mathrm{dom}(g)$ and as $(x, z) \in f$ we have $x \in f^{-1}(\mathrm{dom}(g))$ so

$$\mathrm{dom}(g \circ f) \subseteq \mathrm{dom}(f) \bigcap f^{-1}(\mathrm{dom}(g)) \tag{2.1}$$

If $x \in \mathrm{dom}(f) \bigcap f^{-1}(\mathrm{dom}(g))$ then $x \in \mathrm{dom}(f) \wedge x \in f^{-1}(\mathrm{dom}(g))$ so there exists a y such that $(x, y) \in f$ and $y \in \mathrm{dom}(g)$. As $y \in \mathrm{dom}(g)$ there exists a z such that $(y, z) \in g$ hence $(x, z) \in g \circ f$ or $y \in \mathrm{dom}(g \circ f)$. So

$$\mathrm{dom}(f) \bigcap f^{-1}(\mathrm{dom}(g)) \subseteq \mathrm{dom}(g \circ f) \tag{2.2}$$

If $z \in \mathrm{range}(g \circ f)$ then there exists a x such that $(x, z) \in g \circ f$, so there exists a y such that $(x, y) \in f$ and $(y, z) \in g$. Hence $y \in \mathrm{range}(f)$ and $y \in \mathrm{dom}(g)$ and as $(y, z) \in g$ we conclude that $z \in g(\mathrm{range}(f) \bigcap \mathrm{dom}(g))$ hence

$$\mathrm{range}(g \circ f) \subseteq g\big(\mathrm{range}(f) \bigcap \mathrm{dom}(g)\big) \tag{2.3}$$

If $y \in g(\mathrm{range}(f) \bigcap \mathrm{dom}(g))$ then there exists a $x \in \mathrm{range}(f) \bigcap \mathrm{dom}(g)$ such that $(x, y) \in g$. As $x \in \mathrm{range}(f)$ there exists a z such that $(z, x) \in f$ hence $(z, y) \in g \circ f$ proving that $y \in \mathrm{range}(g \circ f)$. So

$$g\big(\mathrm{range}(f) \bigcap \mathrm{dom}(g)\big) \subseteq \mathrm{range}(g \circ f) \tag{2.4}$$

Using 2.1,2.2 we have that $\mathrm{dom}(g \circ f) = \mathrm{dom}(f) \bigcap f^{-1}(\mathrm{dom}(g))$ and using 2.3,2.4 we have then $g(\mathrm{range}(f) \bigcap \mathrm{dom}(g)) = \mathrm{range}(g \circ f)$ proving the theorem. □

2.1.2 Functions

A function $f: A \to B$ is a partial function so that the mapping exists for every element in A. In other words for every $a \in A$ we can find a unique element b in B.

Definition 2.9. (function) *A partial function $f: A \to B$ is a **function** if* $\forall x \in A \vDash \exists y \in B \vdash (x, y) \in f$

A alternative definition of a function is given by the following theorem.

Theorem 2.10. *A partial function $f: A \to B$ is a function iff* $\mathrm{dom}(f) = A$

Proof.

⇒. Assume that f is a function. If $x \in \mathrm{dom}(f) \Rightarrow \exists y \vdash (x, y) \in f \subseteq A \times B \Rightarrow x \in A \Rightarrow \mathrm{dom}(f) \subseteq A$. If $x \in A \underset{f\,\text{is a function}}{\Rightarrow} \exists y \in B \vdash (x, y) \in f \Rightarrow x \in \mathrm{dom}(f) \Rightarrow A \subseteq \mathrm{dom}(f)$. From this we conclude that $\mathrm{dom}(f) = A$

⇐. Assume that f is a partial function and that $\mathrm{dom}(f) = A$. Then

$$\begin{aligned} x \in A &\Leftrightarrow x \in \mathrm{dom}(f) \\ &\Leftrightarrow \exists y \in B \vdash (x, y) \in f \\ &\Rightarrow \forall x \in A \vDash \exists y \in B \vdash (x, y) \in f \end{aligned}$$

□

Example 2.11. $\emptyset: \emptyset \to \emptyset$ is a function (the null function).

Proof. Using 1.42 we have $\emptyset = \emptyset \times \emptyset \Rightarrow \emptyset \subseteq \emptyset \times \emptyset$ the other requirements for a function are vacuously satisfied. □

If we use the notation $y = f(x)$ iff $(x, y) \in f$ we can rewrite the above theorem as follows

Theorem 2.12. *is a function iff*

1. $\forall x \in A \vDash \exists y \in B \vdash y = f(x)$

2. *If* $y_1 = f(x) \wedge y_2 = f(x) \Rightarrow y_1 = y_2$

Proof.

1. This is equivalent with saying that $f \subseteq A \times B$, $\mathrm{dom}(f) = A$
2. This is equivalent with $(x, y_1), (x, y_2) \in f$ then $y_1 = y_2$ □

Theorem 2.13. *If* $f: A \rightarrow B$ *and* $g: A \rightarrow B$ *are two functions then* $f = g$ *iff* $\forall x \in A \vdash f(x) = g(x)$

Proof.

⇒. Assume that $f = g$ then if $x \in A = \mathrm{dom}(f) = \mathrm{dom}(g)$ there exists a $y, y' \vdash (x, y) \in f \wedge (x, y') \in g \underset{f=g}{\Rightarrow} (x, y), (x, y') \in g \underset{g\,\mathrm{is\,a\,function}}{\Rightarrow} y = y' \Rightarrow f(x) = y = y' = g(x) \Rightarrow f(x) = g(x)$.

⇐. Assume that $\forall x \in A$ we have $f(x) = g(x)$ then

$$\begin{aligned}(x, y) \in f &\Leftrightarrow x \in A \wedge y = f(x)\\ &\Leftrightarrow x \in A \wedge y = g(x)\\ &\Leftrightarrow (x, y) \in g\end{aligned}$$

□

Example 2.14. (Characteristics Function) If A is a class and $B \subseteq A$ then the characteristics function $\mathcal{X}_B: A \rightarrow \{0, 1\}$ is defined as follows $\mathcal{X}_B = (B \times \{1\}) \bigcup ((A \backslash B) \times \{0\})$.

Proof. We use 2.3 to prove that $\mathcal{X}_B$ is a function

1. If $y \in \mathrm{range}(\mathcal{X}_B) \Rightarrow \exists x \vdash (x, y) \in \mathcal{X}_B \Rightarrow x \in A = B \bigcup (A \backslash B)$ and we have two possible cases
   a. $x \in B \Rightarrow y = 1$
   b. $x \notin B \Rightarrow x \in A \backslash B \Rightarrow y = 0$

   so $\mathrm{range}(\mathcal{X}_B) \subseteq \{0, 1\}$
2. If $x \in A$ we have the following cases
   a. $x \in B \Rightarrow (x, 1) \in \mathcal{X}_B$
   b. $x \notin B \Rightarrow x \in (A \backslash B) \Rightarrow (x, 0) \in \mathcal{X}_B$
3. If $(x, y), (x, y') \in \mathcal{X}_B$ then we have two excluding cases
   a. If $x \in B \Rightarrow x \notin (A \backslash B) \Rightarrow (x, y), (x, y') \in B \times \{1\} \Rightarrow y = 1 = y'$
   b. If $x \in A \backslash B \Rightarrow (x, y), (x, y') \in (A \backslash B) \times \{0\} \Rightarrow y = 0 = y'$ □

Definition 2.15. *A partial function* $f: A \rightarrow B$ *is injective iff* $(x, y) \in f \wedge (x', y) \in f \Rightarrow y = y'$. *Using the* $f(x)$ *notation this is equivalent with* $f(x) = f(x') \Rightarrow x = x'$.

Theorem 2.16. *If* $f: A \rightarrow B$ *and* $g: C \rightarrow D$ *are injective partial functions then* $g \circ f: A \rightarrow D$ *is a injective partial function.*

Proof. By 2.6 $g\circ f$ is a partial function. Now if $(x,z),(x',z)\in g\circ f$ then $\exists y,y'\vdash(x,y),(x',y')\in f\wedge(y,z),(y',z)\in g\underset{g\,\text{is injective}}{\Rightarrow}y=y'\Rightarrow(x,y),(x',y)\in f\underset{f\,\text{is injective}}{\Rightarrow}x=x'$ proving that $g\circ f$ is injective. □

Definition 2.17. *A partial function $f: A\rightarrow B$ is invertible if $f^{-1}: B\rightarrow A$ is a partial function*

Lemma 2.18. *A partial function $f: A\rightarrow B$ is invertible if and only if f is injective*

Proof.

- **⇒.** Assume that $f^{-1}: B\rightarrow A$ is a partial function then if $(x,y),(x',y)\in f\Rightarrow(y,x),(y,x')\in f^{-1}\underset{f^{-1}\,\text{is a partial function}}{\Rightarrow}x=x'$ so we have that $f: A\rightarrow B$ is injective.
- **⇐.** Assume now that f is injective then we use 2.3 to prove that f^{-1} is a partial function.
  1. $x\in\text{range}(f^{-1})\Rightarrow\exists y\vdash(y,x)\in f^{-1}\Rightarrow(x,y)\in f\Rightarrow x\in\text{dom}(f)\underset{f\,\text{is a partial function}}{\Rightarrow}x\in A$
  2. $x\in\text{dom}(f^{-1})\Rightarrow\exists y\vdash(x,y)\in f^{-1}\Rightarrow(y,x)\in f\Rightarrow x\in\text{range}(f)\underset{f\,\text{is a partial function}}{\Rightarrow}x\in B$
  3. $(x,y),(x,y')\in f^{-1}\Rightarrow(y,x),(y',x)\in f\underset{f\,\text{is injective}}{\Rightarrow}y=y'$ □

Corollary 2.19. *If $f: A\rightarrow B$ is a injective function then $f^{-1}: f(B)\rightarrow A$ is a function*

Proof. By the above lemma we have that $f^{-1}: B\rightarrow A$ is a partial function. To prove that it is a function, let $x\in f(B)$ then there exist a $y\in B$ such that $y=f(x)$ or $x=f^{-1}(y)$, using 2.12 we have then that $f^{-1}: f(B)\rightarrow A$ is a function. □

Definition 2.20. *A partial function $f: A\rightarrow B$ is surjective iff $\forall y\in B\vDash\exists x\in A\vdash(x,y)\in f$ (or $\forall y\in B\vDash\exists x\in A\vdash y=f(x)$)*

Note 2.21. A partial function $f: A\rightarrow B$ is surjective iff $\text{range}(f)=B$

Proof.

- **⇒.** Assume that f is surjective. If $y\in\text{range}(f)\Rightarrow\exists x\in A\vdash(x,y)\in f\subseteq A\times B\Rightarrow y\in B\Rightarrow\text{range}(f)\subseteq B$. If $y\in B\underset{f\,\text{is surjective}}{\Rightarrow}\exists x\in A\vdash(x,y)\in f\Rightarrow y\in\text{range}(f)\Rightarrow B\subseteq\text{range}(f)$. So we conclude that $\text{range}(f)=B$
- **⇐.** Assume that $\text{range}(f)=B$. Then if $y\in B\Rightarrow y\in\text{range}(f)\Rightarrow\exists x\in A\vdash(x,y)\in f\Rightarrow\forall y\in B\vDash\exists x\in A\vdash(x,y)\in f\Rightarrow f$ is surjective. □

2.1.3 Bijections

Definition 2.22. *A function $f: A\rightarrow B$ is bijective iff f is surjective and f is injective.*

Definition 2.23. *If A and B are classes such that there exists a bijective function $f: A \rightarrow B$ between A and B then we say that A an B are* ***bijective*** *or A and B are in* ***a one to one correspondence*** *and we note this as $A \approx B$*

Example 2.24. (Identity Function) If A is a class then $\mathbb{1}_A: A \rightarrow A$ defined by $\mathbb{1}_A = \{(x, x) | x \in A\}$ is a bijection.

Proof. We have trivially that $\mathbb{1}_A \subseteq A \times A$ and as $(x, y) \in \mathbb{1}_A \wedge (x, y') \in \mathbb{1}_A \Rightarrow y = x \wedge y' = x \Rightarrow y = y'$ $\mathbb{1}_A$ is a partial function. Also $\forall x \in A \vDash (x, x) \in \mathbb{1}_A \Rightarrow \forall x \in A \vDash \exists y \in A \vdash (x, y) \in \mathbb{1}_A \Rightarrow f$ is a function. Further $\forall x \in A \vDash (x, x) \in \mathbb{1}_A \Rightarrow \forall x \in A \vDash \exists y \in A \vdash (y, x) \in \mathbb{1}_A \Rightarrow \mathbb{1}_A$ is surjective. Finally if $(x, y) \in \mathbb{1}_A \wedge (x', y) \in \mathbb{1}_A \Rightarrow x = y \wedge x' = y \Rightarrow x = x'$ so $\mathbb{1}_A$ is injective and thus bijective. □

Example 2.25. (Inclusion Function) If A is a class and $B \subseteq A$ is a subclass then $i_B: B \rightarrow A$ defined by $i_B = \{(x, x) | x \in B\}$ is a injective function.

Proof. We have trivially that $i_B \subseteq B \times B \subseteq B \times A$ and $(x, y) \in i_B \wedge (x, y') \in e_B \Rightarrow y = x = y' \Rightarrow y = y'$ so i_B is a partial function. Next if $x \in B$ then $(x, x) \in i_B \Rightarrow \forall x \in B \vDash \exists y \in A \vdash (x, y) \in i_B \Rightarrow i_B$ is a function. Finally if $(x, y) \in i_B \wedge (x', y) \in \mathrm{ei}_B \Rightarrow x = y = x' \Rightarrow x = x'$ so ei_B is injective. □

Example 2.26. $\emptyset: \emptyset \rightarrow \emptyset$ is a a bijection

Proof.

1. **(injective)** $\forall x, x' \in \emptyset$ we have that $\emptyset(x) = \emptyset(x')$ is satisfied vacuously.
2. **(surjective)** $\forall y \in \emptyset$ we have that there exists a $x \in \emptyset$ is satisfied vacuously. □

2.1.4 Operations on functions and partial functions

Theorem 2.27. *If $f: A \rightarrow B$ is a partial function and $C \subseteq A$ is a subclass of A then $f_{|C}: C \rightarrow B$ with $f_{|C} = \{(x, y) \in f | x \in C\}$ is a partial function with* $\mathrm{dom}(f_{|C}) = \mathrm{dom}(f) \bigcap C$. *Furthermore if f is a function then $f_{|C}$ is a function*

Proof.

1. If $(x, y) \in f_{|C}$ then $(x, y) \in f \subseteq A \times B \wedge x \in C$ so that $(x, y) \in A \times B \wedge x \in C$ or $(x, y) \in C \times B$ proving that $f_{|C} \subseteq C \times B$.
2. If $(x, y), (x, y') \in f_{|C} = \{(x, y) \in f | x \in C\} \subseteq f$ then $(x, y), (x, y') \in f \underset{f \text{ is a partial function}}{\Rightarrow} y = y'$

(1),(2) proves by 2.3 that $f_{|C}: C \rightarrow B$ is a partial function.

For the domain note that

$$\begin{aligned} x \in \mathrm{dom}(f_{|C}) &\Leftrightarrow \exists y \in B \text{ such that } (x, y) \in f_{|C} \\ &\Leftrightarrow \exists y \in B \text{ such that } (x, y) \in f \wedge x \in C \\ &\Leftrightarrow x \in \mathrm{dom}(f) \wedge x \in C \\ &\Leftrightarrow x \in \mathrm{dom}(f) \bigcap C \end{aligned}$$

priving that

$$\mathrm{dom}(f_{|C}) = \mathrm{dom}(f) \bigcap C$$

Finally if f is a function then $\mathrm{dom}(f) = A$ so that $\mathrm{dom}(f_{|C}) = A \bigcap C \underset{C \subseteq A}{=} A$ proving that $f_{|C}: C \rightarrow B$ is a function. □

Theorem 2.28. *If $f: A \rightarrow B$ is a partial function, $C, D \subseteq A$ then $f_{|C \bigcap D} = (f_{|C})_{|D}$. Further if $D \subseteq C$ then $f_{|D} = (f_{|C})_{|D}$*

Proof. We have

$$\begin{aligned}
(x,y) \in f_{|C \bigcap D} &\Rightarrow (x,y) \in f \wedge x \in C \bigcap D \\
&\Rightarrow (x,y) \in f \wedge x \in C \wedge x \in D \\
&\Rightarrow (x,y) \in f_{|C} \wedge x \in D \\
&\Rightarrow (x,y) \in (f_{|C})_{|D} \\
(x,y) \in (f_{|C})_{|D} &\Rightarrow (x,y) \in f_{|C} \wedge x \in D \\
&\Rightarrow (x,y) \in f \wedge x \in C \wedge x \in D \\
&\Rightarrow (x,y) \in f \wedge x \in C \bigcap D \\
&\Rightarrow (x,y) \in f_{|C \bigcap D}
\end{aligned}$$

proving that

$$f_{|C \bigcap D} = (f_{|C})_{|D} \tag{2.5}$$

Further if $D \subseteq C$ then $C \bigcap D = D$ so that $(f_{|C})_{|D} = f_{|C \bigcap D} = f_{|D}$ □

Theorem 2.29. *If $f: A \bigcup B \rightarrow C$ is a partial function then $f = f_{|A} \bigcup f_{|B}$*

Proof. We have

$$\begin{aligned}
(x,y) \in f \subseteq \left(A \bigcup B\right) \times C &\Rightarrow (x \in A \vee x \in B) \wedge (x,y) \in f \\
&\Rightarrow (x \in A \wedge (x,y) \in f) \vee (x \in B \wedge (x,y) \in f) \\
&\Rightarrow (x,y) \in f_{|A} \vee (x,y) \in f_{|B} \\
&\Rightarrow (x,y) \in f_{|A} \bigcup f_{|B} \\
(x,y) \in f_{|A} \bigcup f_{|B} &\underset{f_{|A}, f_{|B} \subseteq f}{\Rightarrow} f \bigcup f = f
\end{aligned}$$

proving that

$$f = f_{|A} \bigcup f_{|B}$$

□

Theorem 2.30. *If $f_1: B \rightarrow A$ and $f_2: C \rightarrow A$ are partial functions with $B \bigcap C = \emptyset$ then if $f = f_1 \bigcup f_2$ the following hold*

1. *$f: B \bigcup C \rightarrow A$ is a partial function and if f_1 and f_2 are functions then $f = f_1 \bigcup f_2: B \bigcup C \rightarrow A$ is a function*
2. *$f_1 = f_{|B}$ and $f_2 = f_{|C}$*

3. *If f_1, f_2 are functions then if $x \in B \Rightarrow f(x) = f_1(x)$ and if $x \in C \Rightarrow f(x) = f_2(x)$*

Proof. We begin the proof by proving

$$(x, y) \in f \wedge x \in B \Leftrightarrow (x, y) \in f_1 \tag{2.6}$$

Proof. If $(x, y) \in f \wedge x \in B$ then $(x, y) \in f_1 \vee (x, y) \in f_2$. Now if $(x, y) \in f_2 \subseteq C \times A \Rightarrow x \in C \Rightarrow x \in B \bigcap C = \emptyset$ a contradiction, so we conclude that $(x, y) \in f_1$. On the other hand if $(x, y) \in f_1 \underset{f_1 \subseteq f \wedge f_1 \subseteq B \times A}{\Rightarrow} (x, y) \in f \wedge x \in B$ □

$$(x, y) \in f \wedge x \in C \Leftrightarrow (x, y) \in f_2 \tag{2.7}$$

Proof. If $(x, y) \in f \wedge x \in C$ then $(x, y) \in f_1 \vee (x, y) \in f_2$. Now if $(x, y) \in f_1 \subseteq B \times A \Rightarrow (x, y) \in B \times A \Rightarrow x \in B \bigcap C = \emptyset$ a contradiction, so we conclude that $(x, y) \in f_2$. On the other hand if $(x, y) \in f_2 \underset{f_2 \subseteq f \wedge f_2 \subseteq C \times A}{\Rightarrow} (x, y) \in f \wedge x \in C$ □

We proceed now as follows

1. If $(x, y) \in f \wedge (x, y') \in f$ then we have the following cases
   a. $x \in B$ then [see 2.6] $(x, y) \in f_1 \wedge (x, y') \in f_1 \underset{f_1 \text{ is } a \text{ function}}{\Rightarrow} y = y'$
   b. $x \in C$ then [see 2.7] $(x, y) \in f_2 \wedge (x, y') \in f_2 \underset{f_2 \text{ is } a \text{ function}}{\Rightarrow} y = y'$.

   As we have also $f = f_1 \bigcup f_2 \subseteq (B \times A) \bigcup (C \times A) \underset{1.46}{=} (B \bigcup C) \times A$ we have that

   $$f: B \bigcup C \rightarrow A \text{ is } a\, a \text{ partial function}$$

   To prove that it is a function consider

   $$\begin{array}{lcl} x \in B \bigcup C & \Rightarrow & x \in B \vee x \in C \\ & \underset{f_1, f_2 \text{ are functions}}{\Rightarrow} & (\exists y \in A \vdash (x, y) \in f_1) \vee (\exists y' \in \vdash (x, y') \in f_2) \\ & \underset{f_1, f_2 \subseteq f_1 \bigcup f_2}{\Rightarrow} & \left(\exists y \in A \vdash (x, y) \in f_1 \bigcup f_2\right) \vee \left(\exists y' \in A \vdash (x, y') \in f_1 \bigcup f_2\right) \\ & \Rightarrow & x \in \operatorname{dom}\left(f_1 \bigcup f_2\right) \vee x \in \operatorname{dom}\left(f_1 \bigcup f_2\right) \\ & \Rightarrow & x \in \operatorname{dom}\left(f_1 \bigcup f_2\right) \end{array}$$

   and thus

   $$f: B \bigcup C \rightarrow A \text{ is } a \text{ function if } f_1, f_2 \text{ are functions}$$

2. We have

   $$\begin{array}{lcl} (x, y) \in f_{|B} & \Leftrightarrow & (x, y) \in f \wedge x \in B \\ & \underset{f_1 \subseteq B \times A \Rightarrow x \in B}{\Leftrightarrow} & (x, y) \in f_1 \\ (x, y) \in f_{|C} & \Leftrightarrow & (x, y) \in f \wedge x \in C \\ & \underset{f_2 \subseteq C \times A \Rightarrow x \in C}{\Leftrightarrow} & (x, y) \in f_2 \end{array}$$

3. Finally we have if $f_1: B \rightarrow A$, $f_2: c \rightarrow A$ are functions so that $f: B \bigcup C \rightarrow A$ is a function that

   a. $x \in B \underset{\mathrm{dom}(f) = B \bigcup C}{\Rightarrow} \exists y \vdash (x, y) \in f \Rightarrow x \in B \wedge (x, y) \in f \Rightarrow (x, y) \in f_1 \Rightarrow f(x) = y = f_1(x)$

   b. $x \in C \underset{\mathrm{dom}(f) = B \bigcup C}{\Rightarrow} \exists y \vdash (x, y) \in f \Rightarrow x \in C \wedge (x, y) \in f \Rightarrow (x, y) \in f_2 \Rightarrow f(x) = y = f_2(x)$

□

Notation 2.31. *If $f_1: B \rightarrow A$ and $f_2: C \rightarrow A$ are (partial) functions, $B \bigcap C = \emptyset$ then $f_1 \bigcup f_2: B \bigcup C \rightarrow A$ is noted as $f: B \bigcup C \rightarrow A$ where $f(x) = \left\{ \begin{smallmatrix} f_1(x) \text{ if } x \in B \\ f_2(x) \text{ if } x \in C \end{smallmatrix} \right.$. This notation can easely be extended to a finite set of functions.*

Theorem 2.32. *If $f: A \rightarrow B$ and $g: C \rightarrow D$ are functions then $g \circ f: A \bigcap f^{-1}(C) \rightarrow D$ is a function.*

Proof. As $f: A \rightarrow B$ and $g: C \rightarrow D$ are functions we have by 2.6 that $g \circ f$ is a partial function. Using 2.8 we have that $\mathrm{dom}(g \circ f) = \mathrm{dom}(f) \bigcap f^{-1}(\mathrm{dom}(g)) \underset{\mathrm{dom}(f) = A \wedge \mathrm{dom}(g) = C \,\mathrm{by}\, 2.10}{=} A \bigcap f^{-1}(C)$ so that by 2.10 again we have that $f: A \bigcap f^{-1}(C) \rightarrow D$ is a function. □

Theorem 2.33. *If $f: A \rightarrow B$ and $g: C \rightarrow D$ are functions with $\mathrm{range}(f) \subseteq C$ then $g \circ f: A \rightarrow D$ is a function.*

Proof. If $\mathrm{range}(f) \subseteq C$ then if $x \in A$ we have $f(x) \in C \Rightarrow f(A) \subseteq C \Rightarrow A \subseteq f^{-1}(C) \Rightarrow A \bigcap f^{-1}(C) = A$. So using the previous theorem we have that $f: A = A \bigcap f^{-1}(C) \rightarrow D$ is a function. □

Theorem 2.34. *We have the following properties of the composition of partial functions*

1. *If $f: A \rightarrow B$, $g: B \rightarrow C$ and $h: C \rightarrow D$ are functions then $h \circ (g \circ f): A \rightarrow D$ is equal to $(h \circ g) \circ f: A \rightarrow E$*
2. *If $f: A \rightarrow B$ is a partial function then $f \circ 1\!\!1_A: A \rightarrow B$ is equal to $f: A \rightarrow B$*
3. *If $f: A \rightarrow B$ is a partial function then $1\!\!1_B \circ f: A \rightarrow B$ is equal to $f: A \rightarrow B$*

Proof.

1.

$$\begin{aligned}
(x, y) \in h \circ (g \circ f) &\Leftrightarrow \exists z \in B \text{ such that } (x, z) \in g \circ f \wedge (z, y) \in h \\
&\Leftrightarrow \exists w \in C, \exists z \in B \text{ such that } (x, w) \in f \wedge (w, z) \in g \wedge (z, \\
&\quad\ y) \in h \\
&\Leftrightarrow \exists w \in C \text{ such that } (x, w) \in f \wedge (w, y) \in h \circ g \\
&\Leftrightarrow (x, y) \in (h \circ g) \circ f
\end{aligned}$$

2.

$$\begin{aligned}(x,y)\in f\circ \mathbb{1}_A &\Leftrightarrow \exists z\in A\,\text{such that}\,(x,z)\in \mathbb{1}_A\wedge (z,y)\in f\\ &\underset{\text{definition of}\,\mathbb{1}_A}{\Leftrightarrow} \exists z\in A\,\text{such that}\,(x,z)=(x,x)\wedge (z,y)\in f\\ &\Leftrightarrow (x,y)\in f\end{aligned}$$

3.

$$\begin{aligned}(x,y)\in \mathbb{1}_B\circ f &\Leftrightarrow \exists z\in B\,\text{such that}\,(x,z)\in f\wedge (z,y)\in \mathbb{1}_B\\ &\underset{\text{definition of}\,\mathbb{1}_B}{\Leftrightarrow} \exists z\in B\,\text{such that}\,(x,z)\in f\wedge (z,y)=(y,y)\\ &\Leftrightarrow (x,y)\in f\end{aligned}$$

□

Theorem 2.35. *If $f: A\rightarrow B$ and $g: C\rightarrow D$ are functions with then $\forall x\in A\bigcap f^{-1}(C)\vDash (g\circ f)(x)=g(f(x))$*

Proof. We use 2.13 so if $x\in A\bigcap f^{-1}(C)$ then

$$\begin{aligned}z=g(f(x)) &\Leftrightarrow (f(x),z)\in g\\ &\Leftrightarrow (y,z)\in g\wedge y=f(x)\\ &\Leftrightarrow (y,z)\in g\wedge (x,y)\in f\\ &\Leftrightarrow (x,z)\in g\circ f\\ &\Leftrightarrow z=(g\circ f)(x)\end{aligned}$$

□

Definition 2.36. *A function $f: A\rightarrow B$ is a invertible if $f^{-1}: B\rightarrow A$ is a function.*

Theorem 2.37. *If $f: A\rightarrow B$ is a function then f is invertible iff f is bijective*

Proof.

- **⇒.** Assume that f is invertible then it is invertible as a partial function and thus by 2.18 we have that f is injective. Now if $x\in B \underset{f^{-1}\,\text{is a function}}{\Rightarrow} \exists y\vdash (x,y)\in f^{-1}\Rightarrow (y,x)\in f\Rightarrow x\in \text{range}(f)\Rightarrow B\subseteq \text{range}(f)\subseteq B$ and f is surjective. So we have that f is bijective.
- **⇐.** Assume that f is bijective then it is injective so that f^{-1} is a partial function by 2.18. Now if $x\in B \underset{f\,\text{is surjective}}{\Rightarrow} \exists y\vdash (y,x)\in f\Rightarrow (x,y)\in f^{-1}\Rightarrow x\in \text{dom}(f^{-1})\Rightarrow B\subseteq \text{dom}(f^{-1})\subseteq B$ so $\text{dom}(f^{-1})=B$ proving that f^{-1} is a function. □

Theorem 2.38. *If $f: A\rightarrow B$ is a bijection then $f^{-1}: B\rightarrow A$ is a bijection.*

Proof. By the previous theorem we have that f^{-1} is invertible and thus a function. Now if $(x,y),(x',y)\in f^{-1}\Rightarrow (y,x),(y,x')\in f \underset{f\,\text{is a partial function}}{\Rightarrow} x=x'\Rightarrow f^{-1}$ is injective. If $x\in A \underset{f\,\text{is a function}}{\Rightarrow} x\in \text{dom}(f)\Rightarrow \exists y\vdash (x,y)\in f\Rightarrow (y,x)\in f^{-1}\Rightarrow x\in \text{range}(f^{-1})\Rightarrow A\subseteq \text{range}(f^{-1})\subseteq A\Rightarrow \text{range}(f^{-1})=A$ and thus f^{-1} is surjective. □

Theorem 2.39. *If $f\colon A\rightarrow B$ is bijective then*

1. $f\circ f^{-1}=\mathbb{1}_B$
2. $f^{-1}\circ f=\mathbb{1}_A$

Proof.

1. $f\circ f^{-1}\colon B\rightarrow B$ is a function as a composition of two functions. Further we have

$$\begin{array}{rcl}
(x,y)\in f\circ f^{-1} & \Rightarrow & \exists z\vdash (x,z)\in f^{-1}\wedge (z,y)\in f\\
 & \Rightarrow & \exists z\vdash (z,x)\in f\wedge (z,y)\in f\\
 & \underset{f\,\mathrm{is}\,a\,\mathrm{function}}{\Rightarrow} & x=y\\
 & \underset{\mathrm{dom}(f\circ f^{-1})=B}{\Rightarrow} & (x,y)\in \mathbb{1}_B\\
(x,y)\in \mathbb{1}_B & \Rightarrow & x\in B\wedge x=y\\
 & \underset{f^{-1}\,\mathrm{is}\,\mathrm{a}\,\mathrm{function}}{\Rightarrow} & \exists z\vdash (x,z)\in f^{-1}\\
 & \Rightarrow & \exists z\vdash (x,z)\in f^{-1}\wedge (z,x)\in f\\
 & \Rightarrow & (x,x)\in f\circ f^{-1}\\
 & \underset{x=y}{\Rightarrow} & (x,y)\in f\circ f^{-1}
\end{array}$$

2. $f^{-1}\circ f\colon A\rightarrow A$ is a function as a composition of two functions. Further we have

$$\begin{array}{rcl}
(x,y)\in f^{-1}\circ f & \Rightarrow & \exists z\vdash (x,z)\in f\wedge (z,y)\in f^{-1}\\
 & \Rightarrow & \exists z\vdash (x,z)\in f\wedge (y,z)\in f\\
 & \underset{f\,\mathrm{is}\,\mathrm{injective}}{\Rightarrow} & x=y\\
 & \underset{\mathrm{dom}(f^{-1}\circ f)=A}{\Rightarrow} & (x,y)\in \mathbb{1}_A\\
(x,y)\in \mathbb{1}_A & \Rightarrow & x\in A\wedge x=y\\
 & \underset{f\,\mathrm{is}\,a\,\mathrm{function}}{\Rightarrow} & \exists z\vdash (x,z)\in f\\
 & \Rightarrow & \exists z\vdash (x,z)\in f\wedge (z,x)\in f^{-1}\\
 & \Rightarrow & (x,x)\in f^{-1}\circ f\\
 & \underset{x=y}{\Rightarrow} & (x,y)\in f^{-1}\circ f
\end{array}$$

$\square$

Theorem 2.40. *If $f\colon A\rightarrow B$ and $g\colon B\rightarrow A$ are two functions such that $f\circ g=\mathbb{1}_B$ and $g\circ f=\mathbb{1}_A$ then f is bijective (hence invertible) and $g=f^{-1}$.*

Proof.

1. **(injectivity)**

$$\begin{array}{rcl}
(x,y),(x',y)\in f & \underset{y\in B,\,g\,\mathrm{is}\,a\,\mathrm{function}}{\Rightarrow} & \exists z\vdash (y,z)\in g\\
 & \Rightarrow & (x,z),(x',z)\in g\circ f=\mathbb{1}_A\\
 & \Rightarrow & x=z=x'\Rightarrow x=x'
\end{array}$$

2. **(surjectivity)**

$$\begin{array}{rcl} y\in B & \Rightarrow & (y,y)\in \mathbb{1}_B = f\circ g \\ & \Rightarrow & \exists z \vdash (y,z)\in g\wedge (z,y)\in f \\ & \Rightarrow & y\in \operatorname{range}(f) \end{array}$$

so we have

$$B\subseteq \operatorname{range}(f)\subseteq B$$

and thus $B = \operatorname{range}(f)$

3. We have

$$\begin{array}{rcl} (x,y)\in g & \underset{g\subseteq B\times A}{\Rightarrow} & y\in A \\ & \underset{\operatorname{dom}(f)=A}{\Rightarrow} & \exists z\vdash (y,z)\in f \\ & \Rightarrow & (x,z)\in f\circ g = 1_B \\ & \Rightarrow & x=z \\ & \Rightarrow & (y,x)\in f \\ & \Rightarrow & (x,y)\in f^{-1} \\ (x,y)\in f^{-1} & \Rightarrow & (y,x)\in f \\ & \underset{f\subseteq A\times B}{\Rightarrow} & x\in B \\ & \underset{\operatorname{dom}(g)=B}{\Rightarrow} & \exists z\vdash (x,z)\in g \\ & \Rightarrow & (y,z)\in g\circ f = 1_A \\ & \Rightarrow & y=z \\ & \Rightarrow & (x,y)\in g \end{array}$$

proving that $g = f^{-1}$ □

We can summarize the two previous theorems in the following

Theorem 2.41. *A function $f: A\to B$ is invertible (or bijective) iff there exists a function $g: B\to A$ such that $f\circ g = \mathbb{1}_B$ and $g\circ f = \mathbb{1}_A$. If such a g exists then $g = f^{-1}$. Note that for $g = f^{-1}$ we have $g\circ f = \mathbb{1}_A\wedge f\circ g = \mathbb{1}_B$ so that $(f^{-1})^{-1} = f$*

Lemma 2.42. *If $f: A\to B$ is a function and $\operatorname{range}(f)\subseteq C$ then $f: A\to C$ is a function*

Proof. We use 2.3 to prove that $f: A\to C$ is a function

1. $\operatorname{range}(f)\subseteq C$ by the assumption of the lemma.
2. $\operatorname{dom}(f) = A$ because $f: A\to B$ is a function.
3. If $(x,y),(x,y')\in f \underset{f:A\to B \text{ is a function}}{\Rightarrow} y = y'$ □

Theorem 2.43. *If $f: A\to B$ and $g: C\to D$ are bijections and $A\bigcap C = \emptyset = B\bigcap D$ then $f\bigcup g: A\bigcup C\to B\bigcup D$ is a bijection.*

Proof. First from the two given functions we form using the previous theorem the functions $f: A \rightarrow B \bigcup D$ and $g: C \rightarrow B \bigcup D$. Using 2.30 and $A \bigcap C = \emptyset$ we have that $f \bigcup g: A \bigcup C \rightarrow B \bigcup D$ is a function. We prove now that it is a bijection

1. **(injectivity)** If $(x, y), (x', y) \in f \bigcup g$ then we have for y the following exclusive possibilities:

   $(x, y) \in f \wedge (x', y) \in f$. then as f is injective we have $x = x'$

   $(x, y) \in f \wedge (x', y) \in g$. then $y \in B \wedge y \in D$ which is impossible as $B \bigcap D = \emptyset$ so this case can not occur.

   $(x, y) \in g \wedge (x', y) \in f$. then $y \in D \wedge y \in B$ which is impossible as $B \bigcap D = \emptyset$ so this case can not occur.

   $(x, y) \in g \wedge (x', y) \in g$. then as g is injective we have $x = x'$

   this proves that in all possible cases we have $x = x'$ proving injectivity.

2. **(surjectivity)** If $y \in B \bigcup D$ then either $y \in B \underset{f \text{ is bijective}}{\Rightarrow} \exists x \in A \vdash (x, y) \in f \subseteq f \bigcup g \Rightarrow (x, y) \in f \bigcup g$ or $y \in D \underset{g \text{ is bijective}}{\Rightarrow} \exists x \in C \vdash (x, y) \in g \subseteq f \bigcup g \Rightarrow (x, y) \in f \bigcup g$ □

Theorem 2.44. *If $f: A \rightarrow B$ is a function, $A \neq \emptyset$ then $f: A \rightarrow B$ is injective if and only if there exists a function $g: B \rightarrow A$ such that $g \circ f = \mathbb{1}_A$*

Proof.

⇒. Suppose f is injective. Take $C = \text{range}(f) = f(A)$ then by the previous lemma we have that $f: A \rightarrow C$ is a function which moreover is surjective (for $C = \text{range}(f)$). Hence f is a bijection. Thus by 2.41 there exists a function $g': f(A) \rightarrow A$ such that $g' \circ f = \mathbb{1}_A$. For $f(A)$ we must consider the following cases:

   $f(A) = B$. then $g = g': B \rightarrow A$ is the required function

   $f(A) \neq B$. as $A \neq \emptyset$ there exists a $x \in A$ then $g: B \rightarrow A$ defined by $g(y) = \begin{cases} g'(y) \text{ if } y \in f(A) \\ x \text{ if } y \in B \setminus f(A) \end{cases}$ (see 2.31) is the desired function.

⇐. Assume that there exists a function $g: B \rightarrow A$ such that $g \circ f = \mathbb{1}_A$ then

$$\begin{array}{rcl} (x, y), (x', y) \in f & \underset{y \in B, \text{dom}(g) = B}{\Rightarrow} & \exists z \vdash (y, z) \in g \\ & \Rightarrow & (x, z), (x', z) \in g \circ f = \mathbb{1}_A \\ & \Rightarrow & x = z = x' \\ & \Rightarrow & x = x' \end{array}$$

□

Theorem 2.45. *A function $f: A \rightarrow B$ is surjective if there exists a $g: B \rightarrow A$ such that $f \circ g = \mathbb{1}_B$.*

Proof.

$$\begin{array}{rcl} x \in B & \Rightarrow & (x, x) \in \mathbb{1}_B = f \circ g \\ & \Rightarrow & \exists z \vdash (x, z) \in g \wedge (z, x) \in f \\ & \Rightarrow & x \in \text{range}(f) \end{array}$$

So we have $B \subseteq \text{range}(f) \subseteq B \Rightarrow \text{range}(f) = B$ proving surjectivity. $\square$

Theorem 2.46. *If $f: A \rightarrow B$ and $g: B \rightarrow C$ are partial functions with* $\text{range}(f) \subseteq \text{dom}(g)$ *then we have for the function $g \circ f: A \rightarrow C$*

1. *If $f: A \rightarrow B$ and $g: B \rightarrow C$ are injective then $g \circ f: A \rightarrow C$ is injective.*
2. *If $f: A \rightarrow B$ and $g: B \rightarrow C$ are surjective then $g \circ f: A \rightarrow C$ is surjective.*
3. *If $f: A \rightarrow B$ and $g: B \rightarrow C$ are bijective **functions** then $g \circ f: A \rightarrow C$ is bijective. Or equivalent if f, g are invertible we have that $g \circ f$ is invertible. Furthermore by 1.54 we have that $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.*

Proof.

1. Let $(x, y), (x', y) \in g \circ f \Rightarrow \exists z, z' \vdash (x, z), (x', z') \in f, (z, y), (z', y) \in g \underset{g \,\text{is injective}}{\Rightarrow} z = z' \Rightarrow (x, z), (x', z) \in f \underset{f \,\text{is injective}}{\Rightarrow} x = x'$
2. If $z \in C \underset{g \,\text{is surjective}}{\Rightarrow} \exists y \vdash (y, z) \in g \underset{g \subseteq B \times C}{\Rightarrow} y \in B \underset{f \,\text{is surjective}}{\Rightarrow} \exists x \vdash (x, y) \in f \Rightarrow (x, z) \in g \circ f \Rightarrow \text{range}(g \circ f) \supseteq C \Rightarrow \text{range}(g \circ f) = C$
3. This follows from (1) and (2) $\square$

Note 2.47. To simplify notation we start from now using the following convention. If $f: A \rightarrow B$ is a partial function then $(x, y) \in f$ is written as $y = f(x)$ meaning that $f(x)$ is the unique element of B such that $(x, f(x)) \in f$. Using this notation we have the following

1. If $g: C \rightarrow D$ is also a partial function then

$$\begin{aligned} y = g(f(x)) &\underset{\text{notation}}{=} (f(x), y) \in g \\ &\underset{\text{notation}}{=} \exists z \in B \text{ such that } z = f(x) \wedge (z, y) \in g \\ &\underset{\text{notation}}{=} \exists z \in B \text{ such that } (x, z) \in f \wedge (z, y) \in g \\ &= (x, y) \in g \circ f \\ &\underset{\text{notation}}{=} y = (g \circ f)(x) \end{aligned}$$

   this can easely extends to more then 2 partial functions.

2. If f is injection so that $f^{-1}: B \rightarrow A$ is a partial function then

$$\begin{aligned} f^{-1}(f(x)) &\underset{(1)}{=} (f^{-1} \circ f)(x) \\ &= \mathbb{1}_B(x) \\ &= x \\ f(f^{-1}(x)) &\underset{(1)}{=} (f \circ f^{-1})(x) \\ &= \mathbb{1}_A(x) \\ &= x \end{aligned}$$

.

From now on we start using the notation $y = f(x)$ instead of $(x, y) \in f$. If f has a inverse then $(x, y) \in f \Rightarrow (y, x) \in f^{-1}$ can be written as $(f^{-1})(f(x)) = x$

2.1.5 Images and preimages of functions and partial functions

Theorem 2.48. *Let $f: A \rightarrow B$, $g: C \rightarrow D$ be partial functions and $E \subseteq A$ then $(g \circ f)_{|E} = (g_{|f(E)}) \circ (f_{|E})$. Further $\mathrm{dom}((g \circ f)_{|E}) = \mathrm{dom}(f) \bigcap f^{-1}(\mathrm{dom}(g)) \bigcap E$. If f, g are functions [so that $\mathrm{dom}(f) = A$ and $\mathrm{dom}(g) = C$] then $\mathrm{dom}((g \circ f)_{|E}) = A \bigcap f^{-1}(C) \bigcap E = f^{-1}(C) \bigcap E$ so that $(g \circ f)_{|E}: f^1(C) \bigcap E \rightarrow D$ is a function*

Proof.

$$\begin{array}{rcl} (x,y) \in (g \circ f)_{|E} & \Leftrightarrow & (x,y) \in g \circ f \wedge x \in E \\ & \Leftrightarrow & \exists z \in B \text{ such that } (x,z) \in f \wedge (z,y) \in g \wedge x \in E \\ & \Leftrightarrow & \exists z \in B \text{ such that } z \in f(E) \wedge (x,z) \in f_{|E} \wedge (z,y) \in g \\ & \Leftrightarrow & \exists z \in B \text{ such that } (x,z) \in f_{|E} \wedge (z,y) \in g_{|f(E)} \\ & \Leftrightarrow & (z,y) \in g_{|g(E)} \circ f_{|E} \end{array}$$

proving that

$$(g \circ f)_{|E} = g_{|f(E)} \circ f_{|E}$$

For the domain note that $\mathrm{dom}((g \circ f)_{|E}) \underset{2.27}{=} \mathrm{dom}(g \circ f) \bigcap E \underset{2.8}{=} \mathrm{dom}(f) \bigcap f^{-1}(\mathrm{dom}(g)) \bigcap E$. □

TODO

Theorem 2.49. *If $f: A \rightarrow B$ is a injective function and $C \subseteq A$ then $f_{|C}: C \rightarrow f(C)$ is a bijection. Further we have that $(f_{|C})^{-1} = (f^{-1})_{|f(C)}$ so that the following bijections are equal $(f_{|C})^{-1}: f(C) \rightarrow C, (f^{-1})_{|f(C)}: f(C) \rightarrow C$*

Proof. First we prove bijectivity

1. **(injectivity)** If $(x,y), (x,y') \in f_{|C} \subseteq f \Rightarrow (x,y), (x,y') \in f \Rightarrow y = y' \Rightarrow f_{|C}$ is injective
2. **(surjectivity)** If $y \in f(C) \Rightarrow \exists x \in C \vdash (x,y) \in f \Rightarrow (x,y) \in f \bigcap (C \times f(C)) \Rightarrow (x,y) \in f_{|C} \Rightarrow f_{|C}$ is surjective.

Next

$$\begin{array}{rcl} (x,y) \in (f_{|C})^{-1} & \Rightarrow & (y,x) \in f_{|C} \\ & \Rightarrow & (y,x) \in f \wedge y \in C \\ & \Rightarrow & (x,y) \in f^{-1} \wedge x \in f(C) \\ & \Rightarrow & (x,y) \in (f^{-1})_{|f(C)} \\ (x,y) \in (f^{-1})_{|f(C)} & \Rightarrow & (x,y) \in f^{-1} \wedge x \in f(C) \\ & \Rightarrow & (x,y) \in f^{-1} \wedge \exists y' \in C \text{ such that } (y',x) \in f \\ & \Rightarrow & (y,x) \in f \wedge \exists x' \in C \text{ such that } (y',x) \in f \\ & \underset{f \text{ is injective} \Rightarrow x = x'}{\Rightarrow} & (y,x) \in f \wedge x \in C \\ & \Rightarrow & (y,x) \in f_{|C} \\ & \Rightarrow & (x,y) \in (f_{|C})^{-1} \end{array}$$

which proves that

$$(f^{-1})_{|f(C)} = (f_{|C})^{-1}$$ □

Theorem 2.50. *Let $f: A \rightarrow B$ be a injective partial function [so that $f^{-1}: B \rightarrow A$ is a partial function] then*

1. *If $C \subseteq B$ then $f^{-1}(C) = (f^{-1})(C)$*
2. *If $C \subseteq A$ then $(f^{-1})^{-1}(C) = f(C)$*

Proof.

1. If $x \in f^{-1}(C)$ then $f(x) \in C$ and thus $x = (f^{-1})(f(x)) \in (f^{-1})(C)$. If $x \in (f^{-1})(C)$ then there exists a $y \in C$ with $x = (f^{-1})(y) \Rightarrow f(x) = f((f^{-1})(y)) = (f \circ f^{-1})(y) = y \in C \Rightarrow x \in f^{-1}(C)$.
2. As $f^{-1}: B \rightarrow A$ is also a bijection take $C \subseteq A$ then using (1) we have $(f^{-1})^{-1}(C) = ((f^{-1})^{-1})(C) = f(C)$ □

Theorem 2.51. *If $f: A \rightarrow B$ is a bijection and $C \subseteq A$ then $f_{|(A \backslash C)}: A \backslash C \rightarrow B \backslash f(C)$ is a bijection*

Proof.

1. **(injective)** If $x, x' \in A \backslash C$ is such that $f_{|(A \backslash C)}(x) = f_{|(A \backslash C)}(x') \Rightarrow f(x) = f(x') \Rightarrow x = x'$.
2. **(surjective)** If $y \in B \backslash f(C)$ then as f is surjective there exists a $x \in A$ such that $f(x) = y$, if now $x \in C \Rightarrow y = f(x) \in f(C)$ contradicting $y \in B \backslash f(C)$ so we have $x \notin C \Rightarrow x \in A \backslash C$ proving surjectivity. □

Theorem 2.52. *If $f: A \rightarrow B$ is a function and $C \subseteq B$ then $f(f^{-1}(C)) \subseteq C$/ Further if f is surjective then $f(f^{-1}(C)) = C$*

Proof. If $x \in f(f^{-1}(C)) \Rightarrow \exists y \in f^{-1}(C)$ so that $x = f(y) \underset{y \in f^{-1}(C) \Rightarrow f(y) \in C}{\Rightarrow} x \in C$. If $x \in C \underset{f \,\mathrm{is surjective}}{\Rightarrow} \exists y \in A$ such that $f(y) = x \in C \Rightarrow y \in f^{-1}(C) \Rightarrow x = f(y) \in f(f^{-1}(C))$ □

Theorem 2.53. *If $f: A \rightarrow B$ is a injective partial function and $C \subseteq A$ then $f^{-1}(f(C)) = C$*

Proof. If $x \in f^{-1}(f(C)) \Rightarrow f(x) \in f(C) \Rightarrow \exists y \in C$ such that $f(x) = f(y) \underset{f \,\mathrm{is injective}}{\Rightarrow} x = y \Rightarrow x \in C$. If $x \in C \Rightarrow f(x) \in f(C) \Rightarrow x \in f^{-1}(f(C))$ □

Theorem 2.54. *If $f: A \rightarrow B$ is a function then we have*

1. *If $C, D \subseteq A$ with $C \subseteq D$ then $f(C) \subseteq f(D)$*
2. *If $C, D \subseteq B$ with $C \subseteq D$ then $f^{-1}(C) \subseteq f^{-1}(D)$*

3. *If* $C \subseteq A \wedge D \subseteq B \vdash f(C) \subseteq D$ *then* $C \subseteq f^{-1}(D)$
4. *If* $C \subseteq B$ *then* $f^{-1}(B \setminus C) = A \setminus f^{-1}(C)$
5. *If* $D \subseteq C \subseteq B$ *then* $f^{-1}(C \setminus D) = f^{-1}(C) \setminus f^{-1}(D)$
6. *If* f *is injective [so that* $f^{-1}\colon f(A) \rightarrow A$ *is a function] and* $C \subseteq A$ *then* $(f^{-1})^{-1}(C) = f(C)$

Proof.

1. If $y \in f(C)$ then $\exists x \in C$ such that $y = f(x) \underset{C \subseteq D}{\Rightarrow} x \in D$ such that $y = f(x) \Rightarrow x \in f(D)$
2. If $x \in f^{-1}(C)$ then $f(x) \in C \underset{C \subseteq D}{\Rightarrow} f(x) \in D \Rightarrow x \in f^{-1}(D)$
3. If $x \in C$ then $f(x) \in f(C) \subseteq D \Rightarrow f(x) \in D \Rightarrow x \in f^{-1}(D)$
4. If $x \in f^{-1}(B \setminus C)$ then $f(x) \in B \setminus C \underset{f^{-1}(B) \subseteq A}{\Rightarrow} x \in A \wedge f(x) \notin C \Rightarrow x \in A \wedge x \notin f^{-1}(C) \Rightarrow x \in A \setminus f^{-1}(C)$. If $x \in A \setminus f^{-1}(C) \Rightarrow x \in A \wedge x \notin f^{-1}(C) \Rightarrow f(x) \in B \wedge f(x) \notin C \Rightarrow f(x) \in B \setminus C \Rightarrow x \in f^{-1}(B \setminus C)$
5. If $x \in f^{-1}(C \setminus D)$ then $x \in A \wedge f(x) \in C \setminus D \Rightarrow x \in A \wedge f(x) \in C \wedge f(x) \notin D \Rightarrow x \in f^{-1}(C) \wedge x \notin f^{-1}(D) \Rightarrow x \in f^{-1}(C) \setminus f^{-1}(D)$. If $x \in f^{-1}(C) \setminus f^{-1}(D)$ then $x \in f^{-1}(C) \wedge x \notin f^{-1}(D) \Rightarrow x \in X \wedge f(x) \in C \wedge f(x) \notin D \Rightarrow x \in X \wedge f(x) \in (C \setminus D) \Rightarrow x \in f^{-1}(C \setminus D)$.
6. If $y \in (f^{-1})^{-1}(C)$ then $(f^{-1})(y) \in C$ so that $y = f(f^{-1}(y)) \in f(C) \Rightarrow (f^{-1})^{-1}(C) \subseteq f(C)$. If $y \in f(C)$ then there exists a $x \in C$ such that $y = f(x) \Rightarrow C \ni x = (f^{-1})(y) \Rightarrow y \in (f^{-1})^{-1}(C)$ □

Theorem 2.55. *Let* A, B, C *be sets and* $f\colon A \rightarrow B$, $g\colon B \rightarrow C$ *functions then if* $D \subseteq C$ *we have that* $(g \circ f)^{-1}(D) = f^{-1}(g^{-1}(D))$

Proof.

$$\begin{aligned} x \in (g \circ f)^{-1}(D) &\Leftrightarrow (g \circ f)(x) \in D \\ &\Leftrightarrow g(f(x)) \in D \\ &\Leftrightarrow f(x) \in g^{-1}(D) \\ &\Leftrightarrow x \in f^{-1}(g^{-1}(D)) \end{aligned}$$

□

Definition 2.56. *If* $f\colon A \rightarrow B$ *is a function and* $\{C_i\}_{i \in I}$ *is a family so that* $\forall i \in I$ *we have* $C_i \subseteq A$ *[in other words* $\{C_i\}_{i \in I}$ *is a family of sub-classes of* A*]. Then we define* $\{f(C_i)\}_{i \in I}$ *to be the family formed by* $\operatorname{graph}(\{f(C_i)\}_{i \in I}) = \{(i, x) | i \in I \wedge x \in f(C_i)\}$.

The following theorem motivates why we use the notation $\{f(C_i)\}_{i \in I}$

Theorem 2.57. *If* $f\colon A \rightarrow B$ *is a function and* $\{C_i\}_{i \in I}$ *is a family of sub-classes of* A *and* $G = \langle \operatorname{graph}(\{f(C_i)\}_{i \in I}), I \rangle$ *then* $\forall i \in I$ *we have that* $G_i = f(C_i)$

Proof. If $i \in I$ then if $x \in G_i \Rightarrow (i, x) \in G \Rightarrow i \in I \wedge x \in f(C_i) \Rightarrow x \in f(C_i) \Rightarrow G_i \subseteq f(C_i)$. If $x \in f(C_i) \underset{i \in I}{\Rightarrow} (i, x) \in G \Rightarrow x \in G_i \Rightarrow f(C_i) \subseteq G_i \Rightarrow f(C_i) = G_i$ □

Once we have defined $\{f(C_i)\}_{i \in I}$ we can prove the following theorem.

Theorem 2.58. *If $f: A \rightarrow B$ is a function, let $\{C_i\}_{i \in I}$ be a non empty family [meaning $I \neq \emptyset$] of sub-classes of A and let $\{D_i\}_{i \in I}$ be a family of sub-classes of B then we have*

1. $f(\bigcup_{i \in I} C_i) = \bigcup_{i \in I} f(C_i)$
2. $f^{-1}(\bigcup_{i \in I} D_i) = \bigcup_{i \in I} f^{-1}(D_i)$
3. $f(\bigcap_{i \in I} C_i) \subseteq \bigcap_{i \in I} f(C_i)$, *if f is injective then* $f(\bigcap_{i \in I} C_i) = \bigcap_{i \in I} f(C_i)$
4. $f^{-1}(\bigcap_{i \in I} D_i) = \bigcap_{i \in I} f^{-1}(D_i)$

Proof.

1.

$$\begin{aligned} y \in f\left(\bigcup_{i \in I} C_i\right) &\Leftrightarrow \exists x \in \bigcup_{i \in I} C_i \vdash y = f(x) \\ &\Leftrightarrow \exists i \in I \vdash \exists x \in C_i \vdash y = f(x) \\ &\Leftrightarrow \exists i \in I \vdash y \in f(C_i) \\ &\Leftrightarrow y \in \bigcup_{i \in I} f(C_i) \end{aligned}$$

2.

$$\begin{aligned} x \in f^{-1}\left(\bigcup_{i \in I} D_i\right) &\Leftrightarrow f(x) \in \bigcup_{i \in I} D_i \\ &\Leftrightarrow \exists i \in I \vdash f(x) \in D_i \\ &\Leftrightarrow \exists i \in I \vdash x \in f^{-1}(D_i) \\ &\Leftrightarrow \bigcup_{i \in I} f^{-1}(D_i) \end{aligned}$$

3.

$$\begin{aligned} y \in f\left(\bigcap_{i \in I} C_i\right) &\Leftrightarrow \exists x \in \bigcap_{i \in I} C_i \vdash y = f(x) \\ &\Rightarrow \forall i \in I \vdash x \in C_i \wedge y = f(x) \\ &\Rightarrow \forall i \in I \vdash y \in f(C_i) \\ &\Rightarrow y \in \bigcap_{i \in I} f(C_i) \end{aligned}$$

proving

$$f\left(\bigcap_{i \in I} C_i\right) \subseteq \bigcap_{i \in I} f(C_i) \tag{2.8}$$

If f is injective , let $y\in\bigcap_{i\in I}f(C_i)$ then $\forall i\in I$ there exists a $x_i\in I$ such that $f(x_i)=y$. As $I\neq\emptyset$ there exists a $j\in I$ such that $f(x_j)=y$. Assume now that $x_j\notin\bigcap_{i\in I}C_i$ then $\exists k\in I$ with $x_j\notin C_k$, from $f(x_j)=y=f(x_k)$ we have as f is injective that $x_j=x_k\in C_k$ a contradiciton. Hence we must have that $x_j\in\bigcap_{i\in I}C_i$ proving that $y=f(x_j)\in f(\bigcap_{i\in I}C_i)$ or $\bigcap_{i\in I}f(C_i)\subseteq f(\bigcap_{i\in I}C_i)$ which togehter with (2.8) proves

$$\text{If } f \text{ is injective then } f\left(\bigcap_{i\in I}C_i\right)=\bigcap_{i\in I}f(C_i)$$

4.

$$\begin{aligned}
y\in f^{-1}\left(\bigcap_{i\in I}C_i\right) &\Leftrightarrow f(y)\in\bigcap_{i\in I}C_i\\
&\Leftrightarrow \forall i\in I\vdash f(y)\in C_i\\
&\Leftrightarrow \forall i\in I\vdash y\in f^{-1}(C_i)\\
&\Leftrightarrow y\in\bigcap f^{-1}(C_i)
\end{aligned}$$

□

Theorem 2.59. *Let $f: A\to B$ be injective then if $X,Y\subseteq A$ we have $f(X\setminus Y)=f(X)\setminus f(Y)$*

Proof. Let $y\in f(X\setminus Y)$ then we have $y=f(x)$ where $x\in X\setminus Y$ then $x\in X$ and $x\notin Y$ so that $y=f(x)\in f(X)$, if now $y\in f(Y)$ then there exists a $z\in Y$ such that $y=f(z)=f(x)\underset{\text{injectivity}}{\Rightarrow} z=x\notin Y$ a contradiction, so we have $y\notin f(Y)$. This proves that

$$f(X\setminus Y)\subseteq f(X)\setminus f(Y). \tag{2.9}$$

If $y\in f(X)\setminus f(Y)$ then $y\in f(X)\wedge y\notin f(Y)$ so there exists a $x\in X$ such that $y=f(x)\wedge y\notin f(Y)$. If now $x\in Y$ then $y=f(x)\in f(Y)$ a contradiction so we must have $x\notin Y$ but then $y\in f(X\setminus Y)$ proving that

$$f(X)\setminus f(Y)\subseteq f(X\setminus Y). \tag{2.10}$$

Finally from (2.9) and (2.10) we have

$$f(X\setminus Y)=f(X)\setminus f(Y)$$

□

Theorem 2.60. *If $f: A\to E$ and $g: C\to E$ be injective functions [so that $f: A\to f(A)$ and $g: C\to g(C)$ are bijections (see 2.49)] then*

1. $\operatorname{dom}(f\circ g^{-1})=g(A\bigcap C)$
2. $f\circ g^{-1}=(f\circ g^{-1})_{|g(A\bigcap C)}=f_{|(A\bigcap C)}\circ(g^{-1})_{|g(A\bigcap C)}=f_{|(A\bigcap C)}\circ(g_{|A\bigcap C})^{-1}$
3. *If $D\subseteq A\bigcap C$ then $(f\circ g^{-1})_{|g(D)}=f_{|D}\circ(g_{|D})^{-1}$ and $(f\circ g^{-1})_{|g(D)}: g(D)\to f(D)$ is a bijection that is the composition of the bijections $(g_{|D})^{-1}: g(D)\to D$ and $f_{|D}: D\to f(D)$. Note that if we take $D=A\bigcap C$ then using (1) we have that $f\circ g^{-1}: g(A\bigcap C)\to f(A\bigcap C)$ is a bijection that is the composition of $(g_{|A\bigcap C})^{-1}: g(A\bigcap C)\to A\bigcap C$ and $f_{|A\bigcap C}: A\bigcap C\to f(A\bigcap C)$.*

Proof.

1. Using 2.8 we have that

$$\begin{aligned}
\operatorname{dom}(f\circ g^{-1}) &= \operatorname{dom}(g^{-1})\bigcap (g^{-1})^{-1}(\operatorname{dom}(f))\\
&\underset{\operatorname{dom}(f)=A}{=} \operatorname{dom}(g^{-1})\bigcap (g^{-1})^{-1}(A)\\
&\underset{(g^{-1}):g(C)\rightarrow C \text{ is bijective}}{=} g(C)\bigcap (g^{-1})^{-1}(A)\\
&\underset{2.50}{=} g(C)\bigcap g(A)\\
&\underset{2.58}{=} g(C\bigcap A
\end{aligned}$$

2. We have

$$\begin{aligned}
f\circ g^{-1} &\underset{(1)}{=} (f\circ g^{-1})_{g(A\bigcap C)}\\
&\underset{2.48}{=} f_{|(g^{-1}(g(A\bigcap C)))}\circ (g^{-1})_{g|(A\bigcap C)}\\
&\underset{\text{g is injective}}{=} f_{|A\bigcap C}\circ (g^{-1})_{|g(A\bigcap C)} && (2.11)\\
&\underset{2.49}{=} f_{|A\bigcap C}\circ (g_{|A\bigcap C})^{-1} && (2.12)
\end{aligned}$$

   (2) is then proved by (2.11) and (2.12)

3. We have

$$\begin{aligned}
(f\circ g^{-1})_{|g(D)} &\underset{(2)}{=} (f_{|(A\bigcap C)}\circ (g^{-1})_{|g(A\bigcap C)})_{g(D)}\\
&\underset{2.48}{=} (f_{|A\bigcap C})_{|(g^{-1})(g(D))} \circ\\
&\qquad ((g^{-1})_{|g(A\bigcap C)})_{|g(D)}\\
&\underset{(g^{-1}):g(C)\rightarrow C \text{ is bijective}}{=} (f_{|A\bigcap C})_{|D}\circ ((g^{-1})_{|g(A\bigcap C)})_{|g(D)}\\
&\underset{2.28}{=} f_{|(A\bigcap C)\bigcap D}\circ (g^{-1})_{|g(A\bigcap C)\bigcap g(D)}\\
&\underset{D\subseteq A\bigcap C}{=} f_{|D}\circ (g^{-1})_{|g(A\bigcap C)\bigcap g(D)}\\
&\underset{2.58}{=} f_{|D}\circ (g^{-1})_{|g((A\bigcap C)\bigcap D)}\\
&\underset{D\subseteq A\bigcap C}{=} f_{|D}\circ (g^{-1})_{|g(D)}\\
&\underset{2.49}{=} f_{|D}\circ (g_{|D})^{-1} && (2.13)
\end{aligned}$$

   First note that by 2.49 and injectivity we have that $g_{|D}: D\rightarrow g(D)$ is a bijection with $(g_{|D})^{-1}: g(D)\rightarrow D$ as it's inverse bijection and that $f_{|D}: D\rightarrow f(D)$ is a bijection. Using 2.46 we have then that the composition of these two functions $(f_{|D})\circ (g_{|D})^{-1}: g(D)\rightarrow f(D)$ is a bijection where by (2.13) $(f\circ g^{-1})_{|g(D)} = f_{|D}\circ (g_{|D})^{-1}$ □

Theorem 2.61. *If $f: A\rightarrow B$ is a function and $C\subseteq B$ then ${}_{C|}f: f^{-1}(C)\rightarrow C$ defined by ${}_{C|}f = f\bigcap (A\times C)$ is a function. Furthermore if f is injective, surjective or bijective then ${}_{C|}f$ is also injective, surjective or bijective*

Proof. First if $(x, y)\in_{C|} f$ then $(x, y)\in f$ and $(x, y)\in A\times C\Rightarrow x\in f^{-1}(C)\Rightarrow (x, y)\in f^{-1}(C)\times C\Rightarrow_{C|} f\subseteq f^{-1}(C)\times C$.

Second if $(x,y),(x,y')\in_{C|} f \Rightarrow (x,y),(x,y')\in f \underset{f\text{ is a function}}{\Rightarrow} y=y'$ so ${}_{C|}f$: $f^{-1}(C)\rightarrow C$ is a partial function. If $x\in f^{-1}(C)$ then there exists a $y\in C$ such that $(x,y)\in f\Rightarrow(x,y)\in f\bigcap(f^{-1}(C)\times C)\Rightarrow(x,y)\in_{C|} f$ so $f^{-1}(C)\subseteq \mathrm{dom}({}_{C|}f)$ and thus ${}_{C|}f$ is a function.

Third if f is injective and $(x,y),(x',y)\in_{C|} f$ then $(x,y),(x',y)\in f\Rightarrow x=x'$ and thus ${}_{C|}f$ is also injective.

Fourth if f is surjective then if $y\in C$ there exists a $x\in A$ such that $(x,y)\in f$ and thus $x\in f^{-1}(C)\Rightarrow(x,y)\in f\bigcap(f^{-1}(C)\times C)={}_{C|}f$ and thus ${}_{C|}f$ is surjective.

Finally if f is bijective then it is injective and surjective and we have just proved that then ${}_{C|}f$ is injective and surjective and thus bijective. □

2.1.6 Constructing families of sets

Theorem 2.62. *If I,A are classes and $f:I\rightarrow\mathcal{P}(A)$ a function. Define then the graph $F=\{(i,x)\in I\times A|x\in f(i)\}\subseteq I\times A$ which has* domain$(F)\subseteq I$ *and defines thus a family $\{F_i\}_{i\in I}$ such that $\forall i\in I$ that $F_i=f(i)$. So we can write $\{F_i\}_{i\in I}=\{f(i)\}_{i\in I}$*

Proof. If $x\in F_i=\{x|(i,x)\in F\}\Rightarrow(i,x)\in F=x\in f(i)\Rightarrow F_i\subseteq f(i)$. If $x\in f(i)\Rightarrow(i,x)\in F\Rightarrow x\in F_i$. □

We show now how we can construct from a family of classes such function

Theorem 2.63. *If $\{F_i\}_{i\in I}$ is a family of classes then there is a graph F with* $\mathrm{dom}(F)=I$. *Define then $f\subseteq I\times\mathcal{P}(\bigcup_{i\in I}F_i)$ by $f=\{(i,A)\in I\times\mathcal{P}(\bigcup_{i\in I}F_i)|A=F_i\}$ then $f:I\rightarrow\mathcal{P}(\bigcup_{i\in I}F_i)$ is a function such that $f(i)=F_i$. In other words $\{F_i\}_{i\in I}=\{f(i)\}_{i\in I}=\{f_i\}_{i\in I}$*

Proof. If $(i,A),(i,A')\in f$ then $A=F_i=A'$ so $f:I\rightarrow\mathcal{P}(\bigcup_{i\in I}F_i)$ is indeed a function. Also $f(i)=A_i$ by definition. □

Theorem 2.64. *If $\{F_i\}_{i\in I}$ is a family of classes defined and $g:J\rightarrow I$ a surjection then by the above theorem we have the existence of a $f:I\rightarrow\mathcal{P}(\bigcup_{i\in I}F_i)$ such that $\{F_i\}_{i\in I}=\{f(i)\}_{i\in I}$. We can form then $f\circ g:J\rightarrow\mathcal{P}(\bigcup_{j\in I}F_i)$ which defines then $\{(f\circ g)(j)\}_{j\in J}=\{f(g(j))\}_{j\in J}$ which is noted by $\{F_{g(j)}\}_{j\in J}$ [here $F_{g(j)}\equiv f(g(j))=(f\circ g)(j)$]. We have then that $\bigcup_{i\in I}F_i=\bigcup_{j\in J}F_{g(j)}$*

Proof. If $x\in\bigcup_{i\in I}F_i$ then $\exists i\in I$ such that $x\in f(i)=F_i \underset{g\text{ is a bijection}}{\Rightarrow}\exists j\in J$ such that $g(j)=i\Rightarrow x\in f(g(j))=F_{g(j)}\Rightarrow x\in\bigcup_{j\in J}F_{g(j)}$. If $x\in\bigcup_{j\in J}F_{g(j)}$ there exists a $j\in J$ such that $x\in F_{g(j)}=f(g(j)) \underset{i=g(j)\in I}{\Rightarrow} x\in f(i)\Rightarrow x\in\bigcup_{i\in I}F_i$ □

2.1.7 Product, Union and Intersection of a family of sets

Axiom 2.65. (Axiom of Replacement) *If A is a set and $f:A\rightarrow B$ is a* ***surjective*** *function then B is a set.*

Corollary 2.66. *If A is a set and $f: A \rightarrow B$ a bijection then B is a set.*

Lemma 2.67. *If $\{A_i\}_{i \in I}$ is a family of sets then $\{A_i | i \in I\} \overset{\text{definition}}{=} \{x | \exists i \in I \vdash x = A_i\}$ (which as A_i is a set and thus a element is a valid definition of a class) then*

$$\begin{aligned}\bigcup_{i \in I} A_i &= \bigcup_{A \in \{A_i | i \in I\}} A \\ \bigcap_{i \in I} A_i &= \bigcap_{A \in \{A_i | i \in I|\}} A\end{aligned}$$

Proof.

$$\begin{aligned}x \in \bigcup_{i \in I} A_i &\Leftrightarrow \exists i \in I \vdash x \in A_i \\ &\Leftrightarrow \exists A \in \{A_i | i \in I\} \vdash x \in A \\ &\Leftrightarrow x \in \bigcup_{A \in \{A_i | i \in I\}} A \\ x \in \bigcap_{i \in I} A_i &\Leftrightarrow \forall i \in I \vdash x \in A_i \\ &\Leftrightarrow \forall A \in \{A_i | i \in I\} \vdash x \in A \\ &\Leftrightarrow \bigcap_{A \in \{A_i | i \in I\}} A\end{aligned}$$

□

Theorem 2.68. *If $\{A_i\}_{i \in I}$ is a family of sets with I a set then $\bigcup_{i \in I} A_i$ is a set.*

Proof. Define $\varphi: I \rightarrow \{A_i | i \in I\}$ by $\varphi = \{(i, A_i) | i \in I\}$ then φ is a surjective function and thus by the axiom of replacement 2.65 we have that $\{A_i | i \in I\}$ is a set. Using then 1.66 we have that $\bigcup_{i \in I} A_i = \bigcup_{A \in \{A_i | i \in I\}} A$ is a set. □

Definition 2.69. *If A and B are sets then the class $B^A = \{f | f \subseteq A \times B \wedge f: A \rightarrow B \text{ is a function}\}$, as $A \times B$ is a set (see 1.79) and thus by the axiom of subsets (see 1.64) we have that f is a set so B^A is indeed a valid class by the axiom of class construction (see 1.9)*

Note 2.70. Using the above definition we see that B^A is the set of function graphs of functions between A and B or in other words $f \in B^A$ is equivalent with $f: A \rightarrow B$ is function. In the future we will often not make a distinction between a function and its function graph.So we will use terms like the function $f \in B^A$ when we actually mean the function graph $f \in B^A$. We will also say that a function $f: A \rightarrow B$ has a certain property P when we actually mean that the function graph has a certain property P. Hence is we define a class (set if B^A is *a* set) $\mathcal{A} = \{f \in B^A | f \text{ satisfies } P\}$ then we call this the class (set) of functions satisfying P.

Theorem 2.71. *Let A, B, C be classes such that $B \subseteq C$ then $B^A \subseteq C^A$*

Proof. Let $f \in B^A$ then $f \subseteq A \times B$ and $\forall x, y \in A$ with $f(x) = f(y)$ we have that $x = y$ which as $A \times B \subseteq A \times C$ proves that $f \in C^A$ □

Theorem 2.72. *If A and B are sets then B^A is a set*

Proof.

If $f\in B^A$ then $f\subseteq A\times B$ and thus $f\in\mathcal{P}(A\times B)$ so that we have $B^A\subseteq\mathcal{P}(A\times B)$. Now by 1.79 we have that $A\times B$ is a set, using the axiom of power sets (see 1.69) we have that $\mathcal{P}(A\times B)$ is a set and finally by the axiom of subsets (see 1.64) we have that B^A is a set. □

Theorem 2.73. *If A is a set then there exists a bijection between 2^A and $\mathcal{P}(A)$*

Proof. Define $\gamma:\mathcal{P}(A)\rightarrow 2^A$ by $\gamma=\{(B,C_B)|B\in\mathcal{P}(A)\}$ (here C_B is the graph of the characteristic function (see 2.14) if $B\in\mathcal{P}_A$ then $C_B=(B\times\{1\}\bigcup((A\setminus B)\times\{0\}))$ then we have that γ is a function by applying 2.3 because:

1. $f\in\text{range}(\gamma)\Rightarrow\exists B\vdash(B,f)\in\gamma\Rightarrow C\in\mathcal{P}(A)\wedge f=C_B\subseteq B\times\{0,1\}=B\times 2\Rightarrow f=C_B: A\rightarrow 2=\{0,1\}$ is *a* function (see 2.14)) $\Rightarrow f\in 2^A$
2. If $B\in\mathcal{P}(A)\Rightarrow B\subseteq A\Rightarrow(B,C_B)\in\gamma\Rightarrow C\in\text{dom}(\gamma)$
3. If $(B,f),(B,f')\in\gamma$ then $f=C_B=f'$

so γ is a function, next we have to prove that it is bijective

1. γ is injective. For if $(D,C_D),(E,C_E)\in\gamma$ with $C_D=C_E$ then we have

$$\begin{aligned} x\in D &\Leftrightarrow C_D(x)=1 \\ &\underset{C_D=C_E}{\Leftrightarrow} C_E(x)=1 \\ &\Leftrightarrow x\in E \end{aligned}$$

   and thus $D=E$

2. γ is surjective. For if $f\in 2^A$ take then $B=\{x\in A|f(x)=1\}$ then we have for $x\in A$

   a. $x\in B\Rightarrow f(x)=1$

   b. $x\notin B\Rightarrow x\in A\setminus B\Rightarrow f(x)\neq 1\underset{\text{range}(f)=\{0,1\}=2}{\Rightarrow}f(x)=0$

   and thus $f=C_B=\gamma(B)$ □

The classical way that a general product of a family of sets $\{A_i\}_{i\in I}$ is as the set of functions from $I\rightarrow\bigcup_{i\in I}A_i$. The problem with this definitions is that subsets of a product of a family of sets can not be itself be a product of a family of $\{B_i\}_{i\in I}$ where $B_i\subseteq A_i$ as this would be functions from $I\rightarrow\bigcup_{i\in I}B_i$ not from $I\rightarrow\bigcup_{i\in I}A_i$. So we would need a kind of function that has no defined destination. This is the idea of a pre-tuple.

Definition 2.74. (pre-tuple) *A pre-tuple $\langle f,I\rangle$ is a pair of a class and a function graph [see 1.60 [so $\forall(x,y),(x,y')\in f$ we have $y=y'$] and $\text{dom}(f)=I$. Note that because of this definition we have that $\forall i\in I$ there exists a unique y such that $(i,y)\in f$ (just like with functions) we use also the same notation to specify this unique value as $f(i)$, essential saying $y=f(i)$ is the same as saying $(i,y)\in f$.*

Example 2.75. $\langle\emptyset,\emptyset\rangle$ is a pretuple for $\forall(x,y),(x,y')\in\emptyset$ we have vacuously that $y=y'$ and $\mathrm{dom}(\emptyset)=\emptyset$

Theorem 2.76. *Let $\langle f,I\rangle$ be a pretuple and let $J\subseteq I$ then if we define $f_{|J}=\{(i,x)\in f|i\in J\}$ then $\langle f_{|J},J\rangle$ is a pretuple. Note that if $(i,x)\in f_{|J}\Rightarrow(i,x)\in f\Rightarrow f_{|J}(i)=f(i)$.*

Proof. If $(i,x),(i,y)\in f_{|J}\Rightarrow(i,x),(i,y)\in f\Rightarrow x=y$ so $f_{|J}$ is a function graph. Also if $i\in\mathrm{dom}(f_{|J})\Rightarrow\exists y$ such that $(i,y)\in f_{|J}\Rightarrow i\in J\Rightarrow\mathrm{dom}(f_{|J})\subseteq J$. If $i\in J\Rightarrow i\in I=\mathrm{dom}(f)\Rightarrow\exists y$ such that $(i,y)\in f\underset{i\in J}{\Rightarrow}(i,y)\in f_{|J}$. □

Now we are ready to define a product of a family of sets.

Definition 2.77. *Let $\{A_i\}_{i\in I}$ be a family of sets where I is also a set (see 1.84) then we define $\prod_{i\in I}A_i=\{f|f\subseteq I\times(\bigcup_{i\in I}A_i)\wedge\langle f,I\rangle\text{ is a pretuple}\wedge\forall i\in I\vDash f(i)\in A_i\}$. Note that $\prod_{i\in I}A_i$ is well defined as a class as by 2.68 $\bigcup_{i\in I}A_i$ is a set so that $I\times(\bigcup_{i\in I}A_i)$ is a set which by the axiom of subsets (see 1.64) gives that f is a set. The elements of $\prod_{i\in\{1,\ldots,n\}}A_i$ are called tuples.*

We prove now that $\prod_{i\in I}A_i$ is also a set

Theorem 2.78. *Let $\{A_i\}_{i\in I}$ be a family of sets where I is also a set (see 1.84) then $\prod_{i\in I}A_i$ is a set*

Proof.

If $f\in\prod_{i\in I}A_i$ then $f\subseteq I\times(\bigcup_{i\in I}A_i)$ so that $f\in\mathcal{P}(I\times(\bigcup_{i\in I}A_i))$, from this it follows that $\prod_{i\in I}A_i\subseteq\mathcal{P}(A\times B)$. Now I is a set, using 2.68 we have that $\bigcup_{i\in I}A_i$ is a set and by 1.79 we have that $I\times(A_i)$ is a set, next using the axiom of power sets (see 1.69) we have that $\mathcal{P}(I\times(\bigcup_{i\in I}A_i))$ is a set and finally by the axiom of subsets (see 1.64) we have that $\prod_{i\in I}A_i$ is a set. □

Example 2.79. If $\{A_i\}_{i\in\emptyset}$ is the empty family of sets then $\bigcup_{i\in\emptyset}A_i=\emptyset$ [if $x\in\bigcup_{i\in\emptyset}A_i$ there exists a $i\in I$ such that $x\in A_i$ which as $I=\emptyset$ is impossible], so in this cases $\prod_{i\in I}A_i=\{\emptyset\}$ (it only contains the empty graph and $\langle\emptyset,\emptyset\rangle$ is a pretuple and of course $\forall i\in I\vdash\emptyset(i)\in A_i$ is satisfied vacuously).

Theorem 2.80. *Let $\{A_i\}_{i\in I}$ be a family of sets where I is also a set then if f: $I\to B$ is a function (f is the graph of our function) such that $\forall i\in I$ we have $f(i)\in A_i$ then $f\in\prod_{i\in I}A_i$*

Proof. As $\forall i\in I$ we have $f(i)\in A_i$ we have that $f\subseteq I\times(\bigcup_{i\in I}A_i)$ and as $f:I\to B$ is a function with domain I we have that $\forall(x,y),(x,y')\in f\Rightarrow y=y'$ and $\mathrm{dom}(f)=i$ so $\langle f,I\rangle$ is also a pretuple, with $\forall i\in I$ that $f(i)\in A_i$ and thus that $f\in\prod_{i\in I}A_i$ □

Theorem 2.81. *Let $\{A_i\}_{i\in I}$ be a family of set, $\beta: I\to J$ a bijection then λ: $\prod_{i\in I}A_i\to\prod_{i\in J}A_{\beta^{-1}(i)}$ defined by $\lambda(x)=x\circ\beta^{-1}$ [which is well defined as β is a bijection] is a bijection*

Proof. First as x, β^{-1} are function graphs and $\operatorname{range}(\beta^{-1}) = I$, $\operatorname{dom}(x) = I$ we have by 1.58 that $\operatorname{dom}(x \circ \beta^{-1}) = I$ so $\langle f, I\rangle$ is a pretuple. As $\forall i \in J$ we have that $x(\beta^{-1}(i)) \in A_{\beta^{-1}(i)}$ we have proved that $x \circ \beta^{-1} \in \prod_{i\in J} A_{\beta^{-1}(i)}$. Further we have

injectivity. if $\lambda(x) = \lambda(y)$ then $x \circ \beta^{-1} = y \circ \beta^{-1} \Rightarrow (x \circ \beta^{-1}) \circ \beta = (y \circ \beta^{-1}) \circ \beta \Rightarrow x = y$

surjectivity. let $y \in \prod_{i\in J} A_{\beta^{-1}(i)}$ and define $x = y \circ \beta$ then as y, β^{-1} are function graphs and $\operatorname{range}(\beta) = I$, $\operatorname{dom}(y) = J$ we have by 1.58 that $\operatorname{dom}(y \circ \beta) = I$. As $\forall i \in I$ we have $y(\beta(i)) \in A_{\beta^{-1}(\beta(i))} = A_i$ we have that $y \circ \beta \in \prod_{i\in I} A_i$. Further we have $\lambda(x) = (y \circ \beta) \circ \beta^{-1} = y$ which proves that λ is surjective.

So λ is a bijection. $\square$

Theorem 2.82. *Let $\{A_i\}_{i\in I}$ be a family of sets where $\forall i \in I$ we have $A_i = A$ then $\prod_{i\in I} A_i = A^I$*

Proof. If $f \in \prod_{i\in I} A_i$ then $f \subseteq I \times \bigcup_{i\in I} A_i = I \times A$, $\langle f, I\rangle$ is a pretuple and $\forall i \in I$ we have $f(i) \in A_i = A$. Then we have

$\operatorname{range}(f) \subseteq A$. If $y \in \operatorname{range}(f)$ then $\exists i \in I$ with $(i, y) \in f \Rightarrow y = f(i) \in A$ proving that $\operatorname{range}(f) \subseteq A$

$\operatorname{dom}(f) = I$. this follows from the fact that $\langle f, I\rangle$ is a pretuple

$\forall (x, y), (x, y') \in f \Rightarrow y = y'$. this follows from the fact that $\langle f, I\rangle$ is a pretuple

So using the definition of a partial function (see 2.3) and a function (see 2.10 we have that $\langle f, I, A\rangle$ is a function which means that by definition of A^I (see 2.69)we have $f \in A^I$. This proves that A $\prod_{i\in I} A_i \subseteq A^I$. Next if $f \in A^I$ then $f \subseteq I \times A = I \times \bigcup_{i\in I} A_i$ and $\langle f, I, A\rangle$ is a function, as also $\forall i \in I \vDash f(i) \in A$ [as $\operatorname{range}(f) \subseteq A$] we have by 2.80 that $f \in \prod_{i\in I} A_i$ proving that $A^I \subseteq \prod_{i\in I} A_i$. So finally we conclude that $A^I = \prod_{i\in I} A_i$. $\square$

Up to now we have the implication $f \in \prod_{i\in I} A_i \Leftrightarrow f \subseteq I \times (\bigcup_{i\in I} A_i) \wedge \langle f, I\rangle$ is a pretuple $\wedge\ \forall i \in I \vDash f(i) \in A_i$, however the last condition implies the first as is expressed in the following theorem.

Theorem 2.83. *Let I be a set, $\{A_i\}_{i\in I}$ be a family of sets then we have $f \in \prod_{i\in I} A_i \Leftrightarrow \langle f, I\rangle$ is a pretuple $\wedge\ \forall i \in I \vDash f(i) \in A_i$*

Proof.

$\Rightarrow$. If $f \in \prod_{i\in I} A_i \Rightarrow f \subseteq I \times (\bigcup_{i\in I} A_i) \wedge \langle f, I\rangle$ is a pretuple $\wedge\ \forall i \in I \vDash f(i) \in A_i \Rightarrow \langle f, I\rangle$ is a pretuple $\wedge\ \forall i \in I \vDash f(i) \in A_i$

$\Leftarrow$. If $\langle f, I\rangle$ is a pretuple $\wedge\ \forall i \in I \vDash f(i) \in A_i$ then from $\langle f, I\rangle$ is a pretuple we have $\operatorname{dom}(f) = I$ and from $\forall i \in I \vDash f(i) \in A_i \Rightarrow \forall i \in I \vDash (i, f(i)) \in I \times A_i \subseteq I \times (\bigcup_{i\in I} A_i) \underset{\operatorname{dom}(f)=I}{\Rightarrow} f \subseteq I \times (\bigcup_{i\in I} A_i)$, proving that $f \subseteq I \times (\bigcup_{i\in I} A_i)$ which together with $\langle f, I\rangle$ is a pretuple $\wedge\ \forall i \in I \vDash f(i) \in A_i$ means that $f \in \prod_{i\in I} A_i$ $\square$

The following notation will be used a lot

Notation 2.84. *Let I be a set, $\{A_i\}_{i\in I}$ be a family of sets and $x\in\prod_{i\in I}A_i$ then given $i\in I$ we note $x(i)$ by x_i so $x_i=x(i)$ and $x_i=y\Leftrightarrow(i,y)\in x$. Using the previous theorem we have then $x\in\prod_{i\in I}A_i\Leftrightarrow\langle x,I\rangle$ is a pretuple and $\forall i\in I$ we have $x_i\in A_i$.*

If $n\in\mathbb{N}$ (see 4.7) then if $I=\{1,\ldots,n\}$ $[\{1,\ldots,n\}=\{x\in\mathbb{N}|1\leqslant x\wedge x\leqslant n\}]$ we use the notation $x=(x_1,\ldots,x_n)$ to mean that $\langle x,I\rangle$ is a pretuple and $\forall i\in\{1,\ldots,n\}$ we have $x(i)=x_i$, using this we have then the equivalence $(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}A_i\Leftrightarrow\forall i\in\{1,\ldots,n\}\vDash x_i\in A_i$, corresponding with the typical mathematical convention for finite products of sets.

Theorem 2.85. *If I is a set and $\{B_i\}_{i\in I}$, $\{A_i\}_{i\in I}$ families of sets such that $\forall i\in I\vDash B_i\subseteq A_i$ then we have $\prod_{i\in I}B_i\subseteq\prod_{i\in I}A_i$.*

Proof. If $x\in\prod_{i\in I}B_i$ then using 2.83 we have $\langle x,I\rangle$ is a pretuple and $\forall i\in I$ we have $x(i)\in A_i\subseteq B_i$ so that $\langle x,I\rangle$ is a pretuple and $\forall i\in I$ we have $x(i)\in A_i$ proving by 2.83 that $x\in\prod_{i\in I}B_i$. □

Lemma 2.86. *Let I be a set and $\{A_i\}_{i\in I}$ a family of sets then $\prod_{i\in I}A_i=\emptyset$ if and only if $\exists i\in I$ with $A_i=\emptyset$*

Proof.

- **⇒.** Let $\prod_{i\in I}A_i=\emptyset$, assume now that $\forall i\in I$ we have $A_i\neq\emptyset$ then using the Axiom of Choice (see 2.201) there exists a function $f: I\rightarrow\bigcup_{i\in I}A_i$ such that $\forall i\in I$ we have $f(i)\in A_i$, so $\langle f,I\rangle$ is a pretuple and as $\forall i\in I$ we have $f(i)\in A_i$ we conclude by 2.83 that $f\in\prod_{i\in I}A_i\Rightarrow f\in\emptyset$ a contradiction. So we must have that $\exists i\in I$ such that $A_i=\emptyset$.
- **⇐.** Let $\exists i\in I$ such that $A_i=\emptyset$ then if $f\in\prod_{i\in I}A_i$ we must have that $f(i)\in A_i=\emptyset$ a contradiction. So we must have that $\prod_{i\in I}A_i=\emptyset$ □

Definition 2.87. *Let $\{A_i\}_{i\in I}$ be a family of sets where I is a set then $\forall i\in I$ we define π_i: $\prod_{i\in I}A_i\rightarrow A_i$ by $\pi_i=\{(x,x(i))|x\in\prod_{i\in I}A_i\}$. We have then that π_i is a function called the projection function. Given $x\in\prod_{i\in I}A_i$ we have $y=\pi_i(x)\Leftrightarrow(x,y)\in\pi_i\underset{x\in\prod_{i\in I}A_i}{\Leftrightarrow}y=x(i)$ so that we have in a shorter notation that $\pi_i(x)=x(i)=x_i$.*

Proof. We still have to prove that π_i is a function, for this we use 2.3

1. $y\in\mathrm{range}(\pi_i)\Rightarrow\exists x\vdash(x,y)\in\pi_i\Rightarrow y=\pi_i(x)=x(i)\in A_i\Rightarrow\mathrm{range}(\pi_i)\subseteq A_i$
2. $x\in\mathrm{dom}(\pi_i)\Rightarrow\exists y\vdash(x,y)\in\pi_i\Rightarrow x\in\Pi_{i\in I}A_i\Rightarrow\mathrm{dom}(\pi_i)\subseteq\Pi_{i\in I}A_i$
3. If $(x,y),(x,y')\in\pi_i$ then $y=\pi_i(x)=x(i)$ and $y'=\pi_i(x)=x(i)\Rightarrow y=y'$

proving that π_i: $\Pi_{j\in I}A_j\rightarrow A_i$ is a partial function. To prove that it is a function consider $x\in\Pi_{j\in I}A_j$ and $i\in\{1,\ldots,n\}$ then $(x,x_i)\in\pi_i\Rightarrow x\in\mathrm{dom}(\pi_i)$ and thus $\Pi_{j\in I}A_j\subseteq\mathrm{dom}(\pi_i)$ proving that π_i: $\prod_{j\in I}A_j\rightarrow A_i$ is a function. □

Theorem 2.88. *Let I be a set and $\{A_i\}_{i\in I}$, $\{B_i\}_{i\in I}$ two families of sets such that $\forall x\in\prod_{i\in I}A_i$ we have that $\forall i\in I$ we have that $\pi_i(x)\in B_i$ then $\prod_{i\in I}A_i\subseteq\prod_{i\in I}B_i$*

Proof. Let $x \in \prod_{i \in I} A_i$ then x is a pre-tuple, further as $\forall i \in I$ we have that $x(i) = \pi_i(x) \in B_i$ so that by 2.80 we have that $x \in \prod_{i \in I} B_i$ $\square$

Theorem 2.89. *Let $\{A_i\}_{i \in I}$ be a family of **non empty** sets where I is a set then $\forall i \in I$ $\pi_i: \prod_{i \in I} A_i \to A_i$ is a surjection.*

Proof. Let $i \in I$ and $x \in A_i$ then as $\forall j \in I \setminus \{i\}$ we have $A_j \neq \emptyset$ we can use the Axiom of Choice (see 2.201) to find a function $f: I \setminus \{i\} \to \bigcup_{j \in I \setminus \{i\}} A_j$ such that $\forall j \in I \setminus \{i\}$ we have $f(j) \in A_j$. Using 2.30 we have then that for $g = f \bigcup \{(i, x)\}$ we have that $g: I \to \bigcup_{i \in I} A_i$ is a function so $\langle g, I \rangle$ is a pretuple and as $\forall j \in I$ we have $g(j) \in A_j$ we conclude by 2.83 that $g \in \prod_{j \in \{1, \ldots, n\}} A_j$. Finally as $g(i) = x$ we have that $(g, x) \in \pi_i$ or $\pi_i(g) = x$ hence π_i is surjective. $\square$

Note 2.90. That the condition that the family of sets is empty is necessary for if $A_1 = \emptyset$ and $A_2 = \{1\}$ then $\prod_{i \in \{1,.2\}} A_i = \emptyset$ and for $1 \in A_2$ we can not find a $x \in \prod_{i \in \{1,.2\}} A_i$ such that $\pi_2(x) = 1$

Theorem 2.91. *Let I be a set and $\{A_i\}_{i \in I}$, $\{B_i\}_{i \in I}$ families of sets with $\forall i \in I$ we have $A_i \neq \emptyset$ (which by 2.86 is equivalent with $\prod_{i \in I} A_i \neq \emptyset$) then $\prod_{i \in I} A_i \subseteq \prod_{i \in I} B_i$ $\Leftrightarrow$ $\forall i \in I$ we have $A_i \subseteq B_i$*

Proof.

- **$\Rightarrow$.** Assume that $\exists i \in I$ with $A_i \nsubseteq B_i$ then there exists a $x \in A_i$ with $x \notin B_i$. Now by the Axiom of Choice (see 2.201) there exists a function $f': I \to \bigcup_{i \in I} A_i$ such that $\forall i \in I \vDash f'(i) \in A_i$, define then $f: I \to \bigcup_{i \in I} A_i$ by $f(j) = \left\{ \begin{array}{l} f(j) \text{ if } j \in I \setminus \{i\} \\ f(j) = x \text{ if } j = i \end{array} \right.$ which is a function (see 2.30) (so $\langle f, I \rangle$ is a pretuple) and as $\forall j \in I$ we have $f(j) \in A_j$ we have by 2.83 that $f \in \prod_{i \in I} A_i$ but as $f(i) = x \notin B_i$ we have that $\prod_{j \in I} A_j \nsubseteq \prod_{j \in I} B_j$ a contradiction. So $\forall i \in I$ we must have $A_i \subseteq B_i$
- **$\Leftarrow$.** This follows from 2.85

$\square$

Note 2.92. The condition that $\forall i \in I$ we have $B_i \neq \emptyset$ is essential in the above theorem. For example if $B_1 = \emptyset, B_2 = \{1, 2\}$ and $A_1 = \{1\}, A_2 = \{2\}$ then clearly $B_1 \subseteq A_1$ but $B_2 \nsubseteq A_2$, still we have $\prod_{i \in \{1,2\}} B_i = \emptyset \subseteq \prod_{i \in \{1,2\}} A_i$.

Corollary 2.93. *Let I be a set then and $\{A_i\}_{i \in I}$, $\{B_i\}_{i \in I}$ be families of sets such that $\forall i \in I$ we have $A_i \neq \emptyset$ then $\prod_{i \in I} A_i = \prod_{i \in I} B_i$ if an only if $\forall i \in I$ we have $A_i = B_i$*

Proof. First as $\forall i \in I$ we have $A_i \neq \emptyset$ so that by 2.86 $\prod_{i \in I} A_i \neq \emptyset$ and thus by 2.86 we have that $\forall i \in I$ we have $B_i \neq \emptyset$.

- **$\Rightarrow$.** If $\prod_{i \in I} A_i = \prod_{i \in I} B_i$ then $\prod_{i \in I} A_i \subseteq \prod_{i \in I} B_i$ and $\prod_{i \in I} B_i \subseteq \prod_{i \in I} A_i \underset{2.91}{\Rightarrow} \forall i \in I$ we have $A_i \subseteq B_i \wedge B_i \subseteq A_i \Rightarrow A_i = B_i$
- **$\Leftarrow$.** If $\forall i \in I$ we have $A_i = B_i \Rightarrow A_i \subseteq B_i \wedge B_i \subseteq A_i \underset{2.85}{\Rightarrow} \prod_{i \in I} A_i \subseteq \prod_{i \in I} B_i \wedge \prod_{i \in I} B_i \subseteq \prod_{i \in I} A_i$ proving that $\prod_{i \in I} A_i = \prod_{i \in I} B_i$ $\square$

Note 2.94. The condition that $\forall i \in I$ we have $B_i \neq \emptyset$ is essential in the above theorem. For example if $B_1 = \emptyset$, $B_2 = \{1, 2\}$ and $A_1 = \{1\}$, $A_2 = \emptyset$ then clearly $A_1 \neq B_1 \wedge A_2 \neq B_2$ and still we have $\prod_{i \in \{1,2\}} B_i = \emptyset = \prod_{i \in \{1,2\}} A_i$.

Theorem 2.95. *If I is a set and $\{A_i\}_{i \in I}$, $\{B_i\}_{i \in I}$ families of sets then $(\prod_{i \in I} A_i) \bigcap (\prod_{i \in I} B_i) = \prod_{i \in I} (A_i \bigcap B_i)$.*

Proof.

1. If $x \in (\prod_{i \in I} A_i) \bigcap (\prod_{i \in I} B_i)$ then using 2.83 $\langle x, I \rangle$ is a pretuple and $\forall i \in I$ we have $x(i) \in A_i \wedge x(i) \in B_i \Rightarrow x(i) \in A_i \bigcap B_i$ so using 2.83 we have $x \in \prod_{i \in I} (A_i \bigcap B_i)$, proving that
$$\left(\prod_{i \in I} A_i\right) \bigcap \left(\prod_{i \in I} B_i\right) \subseteq \prod_{i \in I} (A_i \bigcap B_i) \tag{2.14}$$
.
2. If $x \in \prod_{i \in I} (A_i \bigcap B_i)$ then $\langle x, I \rangle$ is a pretuple and $\forall i \in I$ we have $x(i) \in A_i \bigcap B_i \Rightarrow x(i) \in A_i \wedge x(i) \in B_i$ proving by 2.83 that $x \in \prod_{i \in I} A_i \wedge x \in \prod_{i \in I} B_i \Rightarrow x \in (\prod_{i \in I} A_i) \bigcap (\prod_{i \in I} B_i)$ proving that
$$\prod_{i \in I} (A_i \bigcap B_i) \subseteq \left(\prod_{i \in I} A_i\right) \bigcap \left(\prod_{i \in I} B_i\right) \tag{2.15}$$

From 2.14 and 2.15 it follows then that $(\prod_{i \in I} A_i) \bigcap (\prod_{i \in I} B_i) = \prod_{i \in I} (A_i \bigcap B_i)$ $\square$

We can easily generalize the above theorem to the intersection of a family as is proved in the following theorem.

Theorem 2.96. *Let I, J be sets and $\{\{A_{i,j}\}_{j \in J}\}_{i \in I}$ be a family of families of sets then $\bigcap_{i \in I} (\prod_{j \in J} A_{i,j}) = \prod_{j \in J} (\bigcap_{i \in I} A_{i,j})$*

Proof. First if $x \in \bigcap_{i \in I} (\prod_{j \in J} A_{i,j})$ then $\forall i \in I$ we have $x \in \prod_{i \in j} A_{i,j} \underset{2.83}{\Rightarrow} \langle x, J \rangle$ is a pretuple and $\forall j \in J$ we have $x(j) \in A_j \Rightarrow \langle x, J \rangle$ is a pretuple and $\forall i \in I$, $\forall j \in J \vDash x(j) \in A_{i,j} \Rightarrow \forall j \in J \vDash x(j) \in \bigcap_{i \in I} A_{i,j} \underset{2.83}{\Rightarrow} x \in \prod_{j \in J} (\bigcap_{i \in I} A_{i,j})$ proving that
$$\bigcap_{i \in I} \left(\prod_{j \in J} A_{i,j}\right) \subseteq \prod_{j \in J} \left(\bigcap_{i \in I} A_{i,j}\right) \tag{2.16}$$

Second if $x \in \prod_{j \in J} (\bigcap_{i \in I} A_{i,j})$ then $\langle x, I \rangle$ is a pretuple and $\forall j \in J$ we have $x(j) \in \bigcap_{i \in I} A_{i,j} \Rightarrow \forall i \in I$ we have $x(j) \in A_{i,j}$, giving that $\forall i \in I$ we have $\langle x, I \rangle$ is a pretuple and $\forall j \in J \vDash x(j) \in A_{i,j}$. So $x \in \bigcap_{i \in I} (\prod_{j \in J} A_{i,j})$ proving that
$$\prod_{j \in J} \left(\bigcap_{i \in I} A_{i,j}\right) \subseteq \bigcap_{i \in I} \left(\prod_{j \in J} A_{i,j}\right) \tag{2.17}$$

Using 2.16 and 2.17 proves then $\prod_{j \in J} (\bigcap_{i \in I} A_{i,j}) = \bigcap_{i \in I} (\prod_{j \in J} A_{i,j})$ and the theorem. $\square$

Theorem 2.97. *If I is a set and $\{A_i\}_{i \in I}$, $\{B_i\}_{i \in I}$ families of sets then $(\prod_{i \in I} A_i) \bigcup (\prod_{i \in I} B_i) \subseteq \prod_{i \in I} (A_i \bigcup B_i)$*

Proof. If $x \in (\prod_{i \in I} A_i) \bigcup (\prod_{i \in I} B_i)$ then we have either :

1. $(\boldsymbol{x \in \prod_{i \in I} A_i})$ then $\langle I, x \rangle$ is a pretuple $\wedge$ $\forall i \in I \vDash x(i) \in A_i \subseteq A_i \bigcup B_i$, this means that $x \in \prod_{i \in I} (A_i \bigcup B_i)$.
2. $(\boldsymbol{x \in \prod_{i \in I} B_i})$ then $\langle x, I \rangle$ is a pretuple $\wedge$ $\forall i \in I \vDash x(i) \in B_i \subseteq A_i \bigcup B_i$, this means that $x \in \prod_{i \in I} (A_i \bigcup B_i)$.

so we have always $x \in \prod_{i \in I} (A_i \bigcup B_i)$ and thus $(\prod_{i \in I} A_i) \bigcup (\prod_{i \in I} B_i) \subseteq \prod_{i \in I} (A_i \bigcup B_i)$ □

Note that $(\Pi_{i \in I} A_i) \bigcup (\prod_{i \in I} B_i) \subseteq \prod_{i \in I} (A_i \bigcup B_i)$ but the opposite is not true as the following example shows.

Example 2.98. If $f \in \prod_{i \in \{0,1\}} (A_i \bigcup B_i)$ is such that $f(0) \in A_0$ and $f(1) \in B_1$ then if $A_0 \bigcap B_1 = \emptyset$ we can not have $f \in \prod_{i \in \{0,1\}} A_i$ (because $f(1) \notin A_1$) and we can not have $f \in \prod_{i \in \{0,1\}} B_i$ because $f(0) \in A$

Given a product of sets then there are different representations of the product of sets which are bijective.

Theorem 2.99. *Let I be a set $\{A_i\}_{i \in I}$ a family of sets then if $K, L \subseteq I$ such that $I = K \bigcup L$ and $\emptyset = K \bigcap L$ then $f_{|K} \in \prod_{i \in K} A_i$, $f_{|L} \in \prod_{i \in L} A_i$ and β: $\prod_{i \in I} A_i \rightarrow (\prod_{i \in K} A_i) \times (\prod_{i \in L} A_i)$ defined by $\beta(f) = (f_{|K}, f_{|L})$ is a bijection.*

Proof. First note that for $f \in \prod_{i \in I} A_i$ we have that $\langle I, f \rangle$ is a pretuple and $\forall i \in I$ $f(i) \in A_i$. Using 2.76 it follows that $\langle K, f_{|K} \rangle$ and $\langle L, f_{|L} \rangle$ are pre-tuples and as we have trivially that $\forall i \in K$ $f_{|K}(i) = f(i) \in A_i$ and $\forall i \in L$ $f_{|L}(i) = f(i) \in A_i$ we conclude that $f_{|K} \in \prod_{i \in K} A_i \wedge f_{|L} \in \prod_{i \in L} A_i$. Hence β is indeed a function between $\prod_{i \in I} A_i$ and $(\prod_{i \in K} A_i) \times (\prod_{i \in L} A_i)$. To prove that β is a bijection consider:

injectivity. Let $f, g \in \prod_{i \in I} A_i$ such that $\beta(f) = \beta(g)$ then $(f_{|K}, f_{|L}) = (g_{|K}, g_{|L}) \Rightarrow f_{|K} = g_{|K|} \wedge f_{|L} = g_{|L}$. So if $i \in I$ then we have either:

$\boldsymbol{i \in K}$. then $f(i) = f_{|K}(i) = g_{|K}(i) = g(i)$

$\boldsymbol{i \in L}$. then $f(i) = f_{|L}(i) = g_{|L}(i) = g(i)$

proving that $f = g$

surjection. Let $(f, g) \in (\prod_{i \in K} A_i) \times (\prod_{i \in L} A_i) \Rightarrow f \in \prod_{i \in K} A_i$, $g \in \prod_{i \in L} A_i$ so that f, g are function graphs with domain $\mathrm{dom}(f) = K \wedge \mathrm{dom}(g) = L$. If $h = f \bigcup g$ then if $(x, y), (x, z) \in h = f \bigcup g$ we have either

$\boldsymbol{x \in K \wedge x \notin L}$. then $x \in \mathrm{dom}(f) \wedge x \notin \mathrm{dom}(g)$ so that $(x, y), (x, z) \in f \Rightarrow y = z$

$\boldsymbol{x \notin K \wedge x \in L}$. then $x \notin \mathrm{dom}(f) x \in \mathrm{dom}(g)$ so that $(x, y), (x, z) \in g \Rightarrow y = z$

proving that h is a function graph on $I = K \bigcup L$ and that $h_{|K} = f$ and $h_{|L} = g$. Further $\forall i \in K \bigcup L$ we have either $i \in K \Rightarrow h(i) = f(i) \in A_i$ or $i \in L \Rightarrow L \Rightarrow h(i) = g(i) \in A_i$. This proves that $h \in \prod_{K \bigcup L} A_i$ and that $\beta(h) = (f, g)$ □

2.1.8 Family of elements in a set

Definition 2.100. *If I, A are sets and $x \in A^I$ then we note x as $\{x_i\}_{i \in I}$ where $x_i = x(i)$ and call $\{x_i\}_{i \in I}$ a family of elements of A (or family in A in short). We say that the family is non empty if $I \neq \emptyset$.*

Example 2.101. If A is a set then $1_A: A \to A \in A^A$ defines a family of elements in A which we note by $\{x\}_{x \in A}$ (so $\{x\}_{x \in A} = \{(1_A)_i\}_{i \in A}$) and call the family obtained by self indexing A.

Definition 2.102. *If $\{x_i\}_{i \in I}$ is a family of elements of A (so $\{x_i\}_{i \in I} \in A^I$) and $J \subseteq I$ then $\{x_i\}_{i \in J}$ is defined by $(\{x_i\}_{i \in I})_{|J} \in B^J$.*

Notation 2.103. *If A is a set and $x = \{x_i\}_{i \in I}$ is a family of elements in A, $J \subseteq I$ and $\sigma: J \to J$ a bijection then as x is a function $x: I \to A$ we can consider the function $x \circ \sigma: J \to A$ which represent again a family of elements in A which will be noted as $\{x_{\sigma_i}\}_{i \in J}$. Here $\forall i \in J$ we have $x_{\sigma_i} \equiv (x \circ \sigma)(i) = x(\sigma(i)) \equiv x_{\sigma(i)} = (x \circ \sigma)_i$*

Definition 2.104. *If I, A are sets and $\{x_i\}_{i \in I}$ a family of elements in A then $\{x_i : i \in I\}$ is defined to be equal to $\{x_i\}_{i \in I}(I) = \{y \in A | \exists i \in I \vdash y = \{x_i\}_{i \in I}(i)\} \subseteq A$ (so that $\{x_i : i \in I\}$ must be a set).*

Example 2.105. If A is a set then $\{x\}_{x \in A}(I) = A$ and this is consistent with the notation $A = \{x | x \in A\}$)

Theorem 2.106. *If A, I are sets and $\sigma: I \to I$ a bijection (this is called a permutation) then if $\{x_i\}_{i \in I}$ is a family of elements in A then for the family $\{x_{\sigma_i}\}_{i \in I} (= \{x_i\}_{i \in I} \circ \sigma)$ we have $\{x_i | i \in I\} = \{x_{\sigma_i} | i \in I\}$*

Proof.

$$
\begin{array}{rcl}
y \in \{x_i | i \in I\} & \Rightarrow & \exists i \in I \vdash y = \{x_i\}_{i \in I}(i) \\
& \underset{\sigma \text{ is a bijection}}{\Rightarrow} & \exists j \in I \vdash i = \sigma(j) \\
& \Rightarrow & y = \{x_i\}_{i \in I}(\sigma(j)) \\
& \Rightarrow & y \in \{x_{\sigma_i} | i \in I\} \\
y \in \{x_{\sigma_i} | i \in I\} & \Rightarrow & \exists i \in I \vdash y = (\{x_i\}_{i \in I} \circ \sigma)(i) \\
& \Rightarrow & y = \{x_i\}_{i \in I}(\sigma(i)) \text{ where } \sigma(i) = j \in I \\
& \Rightarrow & y = \{x_i\}_{i \in I}(j) \\
& \Rightarrow & y \in \{x_i | i \in I\}
\end{array}
$$

□

We have now two ways of defining a family of sets that are subsets of a class:

1. If A is a set then $\{A_i\}_{i \in I}$ is a family in $\mathcal{P}(A)$ so $\{A_i\}_{i \in I}$ is a function from I to $\mathcal{P}(A)$
2. If A is a set then $\{A_i\}_{i \in I}$ is a family of sets (see 1.84) such that $\forall i \in I \vDash A_i \subseteq A$ and thus a graph with domain(graph($\{A_i\}_{i \in I}$)) $\subseteq I$

However 2.62 and 2.63 shows that the two definitions (1),(2) are essentially the same if it's about subset's of a certain set.

Definition 2.107. *Let I, A be sets, $B \subseteq A$ and $\{x_i\}_{i \in I}$ a family of elements in A then we say that $\{x_i\}_{i \in I} \subseteq B$ if $\forall i \in I$ we have $\{x_i | i \in I\} \subseteq B$ (or $\{x_i\}_{i \in I}(I) \subseteq B$)*

2.2 Relations

2.2.1 Relations

Definition 2.108. *Let A be a class then a relation in A is any subclass of $A \times A$*

Definition 2.109. *If A is a class then a relation R in A is*

1. **(reflexive)** *iff*
$$\forall x \in A \vdash (x, x) \in R$$
2. **(symmetric)** *iff*
$$(x, y) \in R \;\Rightarrow\; (y, x) \in R$$
3. **(anti-symmetric)** *iff*
$$(x, y) \in R \wedge (y, x) \in R \;\Rightarrow\; x = y$$
4. **(transitive)** *iff*
$$(x, y) \in R \wedge (y, z) \in R \;\Rightarrow\; (x, z) \in R$$

Notation 2.110. *Let A be a class and R is a relation in A then we note $(x, y) \in R$ as* xRy

Definition 2.111. *If A is a class then $I_A = \{(x, x) | x \in A\}$ is the diagonal graph.*

2.2.2 Equivalence relations

Definition 2.112. *If A is a class then a relation in A is a equivalence relation if it is reflexive, symmetric and transitive.*

Definition 2.113. *If A is a set then a partition of A is a family of nonempty subsets of A $\{A_i\}_{i \in I}$ with the following properties*

1. $\bigcup_{i \in I} A_i = A$
2. $\forall i, j \in I$ *we have* $A_i \bigcap A_j = \emptyset \vee A_i = A_j$

Theorem 2.114. *If A is a set and $\{A_i\}_{i \in I}$ is a family of non empty subsets of A then $\{A_i\}_{i \in I}$ is a partition of A iff*

1. $x \in A \Rightarrow \exists i \in I \vdash x \in A_i$
2. $\forall i, j \in I \vdash x \in A_i \bigcap A_j \Rightarrow A_i = A_j$

Proof.

$\Rightarrow$

1. $x \in A = \bigcup_{i \in I} A_i \Rightarrow \exists i \in I \vdash x \in A_i$
2. $\forall i, j \in I$ we have $x \in A_i \bigcap A_j \Rightarrow A_i \bigcap A_j \neq \emptyset \Rightarrow A_i = A_j$

$\Leftarrow$

1. As $\forall i \in I$ we have $A_i \subseteq A \Rightarrow \bigcup_{i \in I} A_i \subseteq A$. If $x \in A \Rightarrow \exists i \in I \vdash x \in A_i \Rightarrow x \in \bigcup_{i \in I} A_i \Rightarrow A \subseteq \bigcup_{i \in I} A_i \Rightarrow A = \bigcup_{i \in I} A_i$
2. $\forall i, j \in I$ if $A_i \bigcap A_j \neq \emptyset \Rightarrow \exists x \in A_i \bigcap A_j \Rightarrow A_i = A_j$ □

Definition 2.115. *If A is a set and R a equivalence relation in A then given a $x \in A$ the equivalence class $R[x]$ of x modulo R is*

$$R[x] = \{y \in A | (y, x) \in R\} \subseteq A$$

(note that because $R[x] \subseteq A$ and A is a set we have by the axiom of subset 1.64 that $R[x]$ is a set)

Theorem 2.116. *If R is a equivalence relation in a set A then*

$$\text{xRy} \Leftrightarrow R[x] = R[y]$$

Proof.

$\Rightarrow$

If xRy then if $z \in R[x] \Rightarrow \text{zRx} \wedge \text{xRy} \Rightarrow \text{zRy} \Rightarrow z \in R[y]$. If $z \in R[y] \Rightarrow \text{zRy} \wedge \text{xRy} \Rightarrow \text{zRy} \wedge \text{yRx} \Rightarrow \text{zRx} \Rightarrow R[x] = R[y]$

$\Leftarrow$

If $R[x] = R[y]$ then as $\text{xRx} \Rightarrow x \in R[x] = R[y] \Rightarrow \text{xRy}$ □

Theorem 2.117. *Let A be a set and R a equivalence relation in A then $\{R[x]\}_{x \in A}$ is a partition of A. $\{R[x]\}_{x \in A}$ is the partition corresponding with R*

Proof. First for every $x \in A$ we have xRx so that $x \in R[x]$ and thus $\{R[x]\}_{x \in A}$ is a family of non empty subsets of A. We use 2.114 to prove that $\{R[x]\}_{x \in A}$ is a partition

1. If $x \in A \Rightarrow \text{xRx} \Rightarrow x \in R[x]$
2. $\forall x, y \in A \vdash z \in R[x] \bigcap R[y] \Rightarrow \text{zRx} \wedge \text{zRy} \Rightarrow \text{xRz} \wedge \text{zRy} \Rightarrow \text{xRy} \underset{\text{previous theorem}}{\Rightarrow} R[x] = R[y]$ □

The above theorem has also a opposite

Theorem 2.118. *If $\{A_i\}_{i \in I}$ is a partition of A then define $R = \{(x, y) | \exists i \in I \vdash x \in A_i \wedge y \in A_i\}$ then R is a equivalence relation, $\forall i \in I \vdash \exists x \in A \vdash A_i = R[x]$ and $\forall x \in A \vdash \exists i \in I \vdash R[x] = A_i$. We call R to be the equivalence relation corresponding with the partition $\{A_i\}_{i \in I}$*

Proof. First we prove that R is a equivalence relation

1. If $x \in A$ then $\exists i \in I \vdash x \in A_i \Rightarrow (x, x) \in R$

2. If $x, y \in A$ with xRy $\Rightarrow \exists i \vdash x \in A_i \wedge y \in A_i \Rightarrow y \in A_i \wedge x \in A_i \Rightarrow (y, x) \in R$
3. If xRy and yRz then $\exists i, j \in I \vdash x \in A_i \wedge y \in A_i \wedge y \in A_j \wedge z \in A_j \underset{y \in A_i \bigcap A_j}{\Rightarrow} A_i = A_j \Rightarrow x \in A_i \wedge z \in A_i \Rightarrow$ xRz

Next if $i \in I$ then as $A_i \neq \emptyset$ there exists a $x \in A_i$ we have then

$$\begin{array}{lll} z \in R[x] & \Leftrightarrow & \mathrm{zRx} \\ & \Leftrightarrow & \exists j \in I \vdash z \in A_j \wedge x \in A_j \\ & \underset{j = i \,\mathrm{as}\, x \in A_i \bigcap A_j}{\Leftrightarrow} & z \in A_i \wedge x \in A_i \\ & \underset{x \in A_i}{\Leftrightarrow} & z \in A_i \end{array}$$

Finally if $x \in A$ then $\exists i \in I$ with $x \in A_i$ and we can use the above again to prove that $A_i = R[x]$. □

Definition 2.119. *If A is set and R is a equivalence relation on A then the set A/R is defined to be $\{R[x] | x \in A\}$ (note that as $R[x] \subseteq A$ we have that $A/R \subseteq \mathcal{P}(A)$ which because of axiom of subsets 1.64 and 1.69 that A/R is a set)*

Theorem 2.120. *If $f: A \rightarrow B$ is a function and R is a equivalence relation on B then $f\langle R\rangle = \{(x, y) | f(x)\mathrm{Rf}(y)\}$ is a equivalence relation. $f\langle R\rangle$ is called the pre-image of R by f*

Proof.

1. **(reflexive)** If $x \in A \Rightarrow f(x) \in B \Rightarrow f(x)\mathrm{Rf}(x) \Rightarrow \mathrm{xf}\langle R\rangle x$
2. **(symmetry)** If $\mathrm{xf}\langle R\rangle y \Rightarrow f(x)\mathrm{Rf}(y) \Rightarrow f(y)\mathrm{Rf}(x) \Rightarrow \mathrm{yf}\langle R\rangle x$
3. **(transitive)** If $\mathrm{xf}\langle R\rangle y \wedge \mathrm{yf}\langle R\rangle z \Rightarrow f(x)\mathrm{Rf}(y) \wedge f(y)\mathrm{Rf}(z) \Rightarrow f(x)\mathrm{Rf}(z) \Rightarrow \mathrm{xf}\langle R\rangle z$ □

Theorem 2.121. *If A is a class with a equivalence relation R then if $B \subseteq A$ then $R_{|B} = \{(x, y) | x \in B \wedge y \in B \wedge (x, y) \in R\}$ is a equivalence relation.*

Proof.

1. **(reflexive)** $x \in B \Rightarrow \mathrm{xRx} \Rightarrow \mathrm{xR}_{|B} x$
2. **(symmetry)** $\mathrm{xR}_{|B} y \Rightarrow x \in B \wedge y \in B \wedge \mathrm{xRy} \Rightarrow x \in B \wedge y \in B \wedge \mathrm{yRx} \Rightarrow \mathrm{yR}_{|B} x$
3. transitivity. If $\mathrm{xR}_{|B} y \wedge \mathrm{yR}_{|B} z \Rightarrow x \in B \wedge y \in B \wedge z \in B \wedge \mathrm{xRy} \wedge \mathrm{yRz} \Rightarrow x \in B \wedge z \in B \wedge \mathrm{xRz} \Rightarrow \mathrm{xR}_{|B} z$ □

Definition 2.122. *Let R, R' be equivalence relations on A then R is a refinement of R' if $R \subseteq R'$ (we express this also by saying that R is finer then R' and that R' is coarser then R*

Theorem 2.123. *If R, R' are equivalence relations on A with $R \subseteq R'$ then we have $z \in R'[x] \Rightarrow R[z] \subseteq R'[x]$*

Proof. If $z \in R'[x]$ then if $y \in R[z] \Rightarrow \mathrm{yRz} \Rightarrow \mathrm{yR}'z \underset{z \in R'[x]}{\Rightarrow} \mathrm{yR}'z \wedge \mathrm{zR}'x \Rightarrow \mathrm{yR}'x \Rightarrow y \in R'[x]$ □

Corollary 2.124. *If R, R' are equivalence relations on A with $R \subseteq R'$ then $\forall x \in A \vdash R[x] \subseteq R'[x]$*

Proof. If $x \in A$ then as $x \in R'[x] \underset{\text{previous theorem}}{\Rightarrow} R[x] \subseteq R'[x]$ □

Corollary 2.125. *If P, R are equivalence relations on a set A with $P \subseteq R$. Then the quotient of $\frac{R}{P}$ is a relation in A/P defined as follows*

$$\frac{R}{P} = \{(P[x], P[y]) | (x, y) \in R\}$$

We have then that $\frac{R}{P}$ is a equivalence relation on A/P

Proof. This is proved by

1. **(reflexive)** If $P[x] \in A/P$ then from $(x, x) \in R$ we have $(P[x], P[x]) \in \frac{R}{P}$
2. **(symmetry)** If $P[x], P[y] \in A/P$ then if $(P[x], P[y]) \in \frac{R}{P} \Rightarrow (x, y) \in R \Rightarrow (y, x) \in R \Rightarrow (P[y], P[x]) \in \frac{R}{P}$
3. **(transitive)** If $(P[x], P[y]) \in \frac{R}{P} \wedge (P[y], P[z]) \in \frac{R}{P}$ then $(x, y) \in R \wedge (y, z) \in R \Rightarrow (x, z) \in R \Rightarrow (P[x], P[z]) \in \frac{R}{P}$ □

2.2.3 Equivalence relations and functions

Theorem 2.126. *If $f: A \rightarrow B$ is a function then $R_f \subseteq A \times A$ defined by $R_f = \{(x, y) \in A \times A | f(x) = f(y)\}$ is a equivalence relation in A, it is called the equivalence relation determined by f*

Proof.

1. **(reflexivity)** If $x \in A \Rightarrow f(x) = f(x) \Rightarrow (x, x) \in R_f$
2. **(symmetry)** If $\mathrm{xR}_f y \Rightarrow f(x) = f(y) \Rightarrow f(y) = f(x) \Rightarrow \mathrm{yR}_f x$
3. **(transitive)** If $\mathrm{xR}_f y$ and $\mathrm{yR}_f z$ then $f(x) = f(y) \wedge f(y) = f(z) \Rightarrow f(x) = f(z) \Rightarrow \mathrm{xR}_f z$ □

So with a function is associated a equivalence relation, we can also do the opposite, given a equivalence relation then associate this equivalence relation with a function.

Theorem 2.127. *If R is a equivalence relation in a set A then $f_R: A \rightarrow A/R$ defined by $f_R = \{(x, R[x]) | x \in A\}$ is a function and we have then $R_{f_R} = R$. Furthermore we have that f_R is surjective. We call f_R the canonical function of R*

Proof. First we use 2.3 to prove that f_R is a function

1. $\mathrm{range}(f_R) \subseteq A/R$. If $y \in \mathrm{range}(f_R) \Rightarrow \exists x \vdash y = R[x] \in A/R \Rightarrow \mathrm{range}(f_R) \subseteq A/R$
2. $\mathrm{dom}(f_R) = A$ If $x \in \mathrm{dom}(f_R) \Rightarrow \exists y \vdash (x, y) \in f_R \Rightarrow x \in A \wedge y = R[x] \Rightarrow x \in A \Rightarrow \mathrm{dom}(f_R) \subseteq A$. If $x \in A \Rightarrow (x, R[x]) \in f_R \Rightarrow x \in \mathrm{dom}(f) \Rightarrow A \subseteq \mathrm{dom}(f)$

3. If $(x, y), (x, y') \in f \Rightarrow y = y'$. If $(x, y), (x, y') \in f_R$ then $x \in A \wedge y = R[x] \wedge y' = R[x] \Rightarrow y = y'$

This proves that $f_R: A \rightarrow A/R$ is a function. Next

$$\begin{aligned}(x, y) \in R &\underset{2.116}{\Leftrightarrow} R[x] = R[y]\\ &\Leftrightarrow f_R(x) = f_R(y)\\ &\Leftrightarrow (x, y) \in R_{f_R}\end{aligned}$$

proving that $R = R_{f_R}$.

Finally we must prove that f_R is surjective, so if $y \in A/R \Rightarrow \exists x \vdash y = R[x] \Rightarrow \exists x \vdash y = f_R(x) = R[x] \Rightarrow y \in \text{range}(f) \Rightarrow \text{range}(f) = A/R$ □

Theorem 2.128. *If $f: A \rightarrow B$ is a function where A, B are sets, define then s_f: $A/R_f \rightarrow f(A)$ with $s_f = \{(R_f[x], f(x)) | x \in A\}$ then we have $f = e_{f(A)} \circ s_f \circ f_{R_f}$ where*

1. *$f_{R_f}: A \rightarrow A/R_f$ is defined in 2.127 and is proved there to be surjective*
2. *$e_{f(A)}: f(A) \rightarrow B$ is the inclusion function (see 2.25) which is injective (see 2.25)*
3. *s_f is bijective*

This is called the canonical decomposition of a function. So we can consider every function as the composition of a surjective, bijective and injective function

Proof. First we prove that s_f is a function. Again we use 2.3

1. $\text{range}(s_f) \subseteq f(A)$. If $y \in \text{range}(s_f) \Rightarrow \exists x \vdash (x, y) \in s_f \Rightarrow \exists z \in A \vdash x = R_f[z]$, $y = f(z) \Rightarrow y \in f(A) \Rightarrow \text{range}(s_f) \subseteq f(A)$
2. $\text{dom}(s_f) = A/R_f$. If $x \in \text{dom}(s_f) \Rightarrow \exists y \vdash (x, y) \in s_f \Rightarrow \exists z \in A \vdash x = R_f[z] \wedge y = f(z) \Rightarrow x \in A/R_f \Rightarrow \text{dom}(s_f) \subseteq A/R_f$. Now if $x \in A/R_f \Rightarrow \exists z \in A \vdash x = R_f[z] \Rightarrow (R_f[z], f(z)) \in s_f \Rightarrow (x, f(z)) \in s_f \Rightarrow x \in \text{dom}(s_f) \Rightarrow A/R_f \subseteq \text{dom}(s_f)$ so we conclude that $\text{dom}(s_f) = A/R_f$
3. If $(x, y), (x, y') \in s_f \Rightarrow y = y'$. This is the most important one to prove, so let $(x, y), (x, y') \in s_f$ then there exists a $z, z' \in A \vdash x = R_f[z] \wedge y = f(z) \wedge x = R_f[z'] \wedge y' = f(z')$ now from $R_f[z] = x = R_f[z'] \Rightarrow R_f[z] = R_f[z']$ we have by 2.116 that $z\mathrm{R}_f z' \Rightarrow f(z) = f(z') \Rightarrow y = y'$

Next we must prove that s_f is a bijection

1. **(injective)** If $(x, y), (x', y) \in s_f$ then $\exists z, z' \in A \vdash x = R_f[z] \wedge x' = R_f[z'] \wedge f(z) = y = f(z') \Rightarrow z\mathrm{R}_f z' \Rightarrow R_f[z] = R_f[z'] \Rightarrow x = x'$
2. **(surjective)** If $y \in f(A) \Rightarrow \exists z \in A \vdash y = f(z) \Rightarrow (R_f[z], f(z)) = (R_f[z], y) \in s_f \Rightarrow y \in \text{range}(s_f) \Rightarrow f(A) \subseteq \text{range}(s_f)$ proving surjectivity.

Finally we must prove that $f = e_{f(A)} \circ s_f \circ f_{R_f}$ now if $x \in A$ then we have by 2.35 that

$$\begin{aligned}(e_{f(A)} \circ s_f \circ f_{R_f})(x) &= (e_{f(A)} \circ s_f)(f_{R_f}(x))\\ &= e_{f(A)}(s_f(f_{R_f}(x)))\\ &= e_{f(A)}(s_f(R_f[x]))\\ &= e_{f(A)}(f(x))\\ &= f(x)\end{aligned}$$

by 2.13 we have then that $f = e_{f(A)} \circ s_f \circ f_{R_f}$

□

So we conclude that $A/R_f \approx f(A)$. Further if f is surjective then $f(A) = \mathrm{range}(f) = B$ and then $e_{f(A)}$ is $i_{f(A)}$ which is a bijection so that by 2.46 we have that $e_{f(A)} \circ s_f: A/R_f \to B$ is a bijection so $A/R_f \approx B$.

Theorem 2.129. *Let $f: A \to B$ be a function between sets A and B. Let $R \subseteq R_f$ be a equivalence relation in A which is finer then R_f then we can define f/R: $A/R \to B$ by $f/R = \{(R[x], f(x)) | x \in A\}$. Then we have that f is a function (called the **quotient of f by R**). Furthermore we have that $\frac{R_f}{R} = R_{f/R}$*

Proof. First we prove that $f/R: A/R \to B$ is a function (using 2.3):

1. $\mathrm{range}(f/R) \subseteq B$. If $y \in \mathrm{range}(f/R) \vdash \exists x \vdash (x, y) \in f/R \Rightarrow \exists z \in A \vdash x = R[x] \wedge y = f(z) \in B \Rightarrow \mathrm{range}(f) \subseteq B$
2. $\mathrm{dom}(f/R) = A/R$. If $x \in \mathrm{dom}(f/R)$ then $\exists y \vdash (x, y) \in f/R \Rightarrow \exists z \in A \vdash x = R[z] \wedge y = f(z) \Rightarrow x \in A/R \Rightarrow \mathrm{dom}(f/R) \subseteq A/R$. If $x \in A/R \Rightarrow \exists z \vdash x = R[z] \Rightarrow (R[z], f(z)) \in f/R \Rightarrow x = R[z] \in \mathrm{dom}(f/R) \Rightarrow A/R \subseteq \mathrm{dom}(f/R)$
3. $(x, y), (x, y') \in f/R \Rightarrow y = y'$. So if $(x, y), (x, y') \in f/R \Rightarrow \exists z, z' \in A \vdash x = R[z] \wedge y = f(z) \wedge x = R[z'] \wedge y' = f(z') \Rightarrow (x, z) \in R \wedge (x, z') \in R \wedge y = f(z) \wedge y = f(z') \underset{R \subseteq R_f}{\Rightarrow} (x, z) \in R_f \wedge (x, z') \in R_f \wedge y = f(z) \wedge y' = f(z) \Rightarrow f(x) = f(z) \wedge f(x) = f(z') \wedge y = f(z) \wedge y' = f(z') \Rightarrow y = f(z) = f(x) = f(z') = y' \Rightarrow y = y'$. Note that $R \subseteq R_f$ is essential to prove that R_f is a function.

Now for the last part

$$\begin{aligned}
(x, y) \in R_{f/R} &\Leftrightarrow x \in A/R \wedge y \in A/R \wedge (f/R)(x) = (f/R)(y) \\
&\Leftrightarrow \exists z, z' \in A \wedge x = R[z] \wedge y = R[z'] \wedge (f/R)(R[z]) = (f/R)(R[z']) \\
&\Leftrightarrow \exists z, z' \in A \wedge x = R[z] \wedge y = R[z'] \wedge f(z) = f(z') \\
&\Leftrightarrow \exists z, z' \in A \wedge x \in R[z] \wedge y = R[z'] \wedge (z, z') \in R_f \\
&\Leftrightarrow (x, y) \in \frac{R_f}{R}
\end{aligned}$$

□

2.3 Partially ordered classes

2.3.1 Order relation

Definition 2.130. (Pre-order) *If A* is *a* class then *a* relation *R in A is a pre-order if it is reflexive and transitive or in other words*

1. $\forall x \in A \vdash \mathrm{xRx}$ *(reflectivity)*

2. $\forall x, y \in A \vdash \mathrm{xRy} \wedge \mathrm{yRz} \Rightarrow \mathrm{xRz}$ *(transitivity)*

Definition 2.131. *A class A together with a pre-order R on the class is called a pre-ordered class and noted by $\langle A, R\rangle$ (if A is a set then $\langle A, R\rangle$ is called a pre-ordered set)*

We can extend the concept of a pre-order to a order using the following definition.

Definition 2.132. (Order relation) *If A is a class then a relation R in A is a order relation if it is reflexive, anti-symmetric and transitive or in other words*

1. $\forall x \in A \vdash \mathrm{xRx}$ *(reflexivity)*
2. $\mathrm{xRy} \wedge \mathrm{yRx} \Rightarrow x = y$ *(anti-symmetry)*
3. $\mathrm{xRy} \wedge \mathrm{yRz} \Rightarrow \mathrm{xRz}$ *(transitivity)*

A pair $\langle A, R\rangle$ where A is a clas and R a order relation is called a ***partially ordered class****.*

Example 2.133. If $\mathcal{A}$ is a class of classes then $\langle A, \subseteq\rangle$ is a partially ordered class

Proof.

1. **(reflexive)** If $A \in \mathcal{A}$ then $A \subseteq A$
2. **(symmetry)** If $A \subseteq B$ and $B \subseteq A$ then $A = B$
3. **(transitive)** If $A \subseteq B$ and $B \subseteq C$ then $A \subseteq C$ □

Example 2.134. Let A be a class, $\langle B, \leqslant\rangle$ a partially ordered class then $\langle B^A, \leqslant\rangle$ where $\leqslant = \{(f, g) \in B^A | \forall x \in A \text{ we have } f(x) \leqslant g(x)\}$ forms a partially ordered class. Here B^A is a the class of all functions from A to B. Note that we say $f \leqslant g$ if $\forall x \in A$ we have $f(x) \leqslant g(x)$

Proof.

1. **(reflective)** Take $f \in B^A$, as $\forall x \in A$ we have $f(x) \in B \underset{\langle B, \leqslant\rangle \text{ is a partial ordered set}}{\Rightarrow} f(x) \leqslant f(x)$ we have $f \leqslant f$
2. **(symmetry)** If $f, g \in B^A$ such that $f \leqslant g \wedge g \leqslant f$ then $\forall x \in A$ we have $f(x) \leqslant g(x) \wedge g(x) \leqslant f(x) \Rightarrow f(x) = g(x) \Rightarrow f = g$.
3. **(transitive)** Let $f, g, h \in B^A$ such that $f < g \wedge g < h$ then $\forall x \in A$ we have $f(x) \leqslant g(x) \wedge g(x) \leqslant h(x) \Rightarrow f(x) \leqslant h(x)$ proving that $f \leqslant h$ □

Definition 2.135. *If we have a class A together with a order relation R in A then the pair $\langle A, R\rangle$ (see 1.80) is called a partially ordered class. If A is a set then $\langle A, R\rangle$ is called a partially ordered set.*

Notation 2.136. *If $\langle A, R\rangle$ is a partially ordered class (set) then R is noted as $\leqslant$ so $x \leqslant y$ is the same as $(x, y) \in R$. In words we say that x is smaller then y. Also $y \geqslant x$ (y is greater then x) is another way of saying $x \leqslant y$. Finally $x < y$ (x is strictly less then y) is a abbreviation for $x \leqslant y \wedge x \neq y$ and $y > x$ (y is strictly greater then x) means $x < y \wedge x \neq y$.*

Theorem 2.137. *If $\langle A, \leqslant\rangle$ is a partially ordered set then*

1. $x \leqslant y \wedge y < z \Rightarrow x < z$
2. $x < y \wedge y \leqslant z \Rightarrow x < z$
3. $x < y \wedge y < z \Rightarrow x < z$
4. $(x < y \vee x = y) \Leftrightarrow (x \leqslant y)$

Proof.

1. $x \leqslant y \wedge y < z \Rightarrow x \leqslant y \wedge y \leqslant z \wedge y \neq z \Rightarrow x \leqslant z \wedge y \neq z$ if now $x = z \Rightarrow z \leqslant y \Rightarrow z = y$ a contradiction so we must have $x \leqslant z \wedge x \neq z \Rightarrow x < z$
2. $x < y \wedge y \leqslant z \Rightarrow x \leqslant y \wedge y \leqslant z \wedge x \neq y \Rightarrow x \leqslant z \wedge x \neq y$ then if $x = z \Rightarrow y \leqslant x \Rightarrow$ $x = y$ a contradiction so we must have $x \leqslant z \wedge x \neq z \Rightarrow x < z$
3. $x < y \wedge y < z \Rightarrow x \leqslant y \wedge y < z \underset{(1)}{\Rightarrow} x < z$
4. We have

$$
\begin{array}{rcl}
(x < y \vee x = y) & \Leftrightarrow & ((x \leqslant y \wedge x \neq y) \vee x = y) \\
& \Leftrightarrow & ((x \leqslant y \vee x = y) \wedge (x \neq y \vee x = y)) \\
& \Leftrightarrow & x \leqslant y \vee x = y \\
& \underset{y \leqslant y \wedge x = y \Rightarrow x \leqslant y}{\Rightarrow} & x \leqslant y \vee x \leqslant y \\
& \Rightarrow & x \leqslant y \\
x \leqslant y & \Rightarrow & x \leqslant y \wedge (x = y \vee x \neq y) \\
& \Rightarrow & (x \leqslant y \wedge x = y) \vee (x \leqslant y \wedge x \neq y) \\
& \Rightarrow & (x = y) \vee (x \leqslant y \wedge x \neq y) \\
& \Rightarrow & x = y \vee x < y
\end{array}
$$

□

Theorem 2.138. *If $\langle A, \leqslant\rangle$ is a pre-ordered class then*

1. *$\sim \in A \times A$ defined by $x \sim y$ if and only if $x \leqslant y \wedge y \leqslant x$ is a equivalence relation.*
2. *If $\sim[x], \sim[y] \in A/\sim$ with $\sim[x] \neq \sim[y]$ then we define $\sim[x] < \sim[y]$ iff $x < y$, this definition is well defined (so independent of the choice of the representation).*
3. *If we define $\sim[x] \leqslant \sim[y]$ by $\sim[x] < \sim[y] \vee \sim[x] = \sim[y]$ then $\langle A/\sim, \leqslant\rangle$ is a ordered class.*
4. *If $x \leqslant y \Rightarrow \sim[x] \leqslant \sim[y]$*
5. *If $\sim[x] \leqslant \sim[y]$ and $x' \in \sim[x], y' \in \sim[y]$ then $x' \leqslant y'$*

Proof.

1. To prove that $\sim$ is a equivalence relation note that
   a. If $x \in A$ then $x \leqslant x \wedge x \leqslant x \Rightarrow x \sim x$
   b. If $x, y \in A$ and $x \sim y$ then $x \leqslant y \wedge y \leqslant x \Rightarrow y \leqslant x \wedge x \leqslant y \Rightarrow y \sim x$
   c. If $x, y, z \in A$ and $x \sim y \wedge y \sim z \Rightarrow x \leqslant y \wedge y \leqslant x \wedge y \leqslant z \wedge z \leqslant y \Rightarrow x \leqslant z \wedge z \leqslant x \Rightarrow x \sim z$

2. If $\sim[x'] = \sim[x] \neq \sim[y] = \sim[y']$ with $y' \leqslant x'$ then $y' \leqslant x' \wedge x' \leqslant x \wedge x \leqslant x' \wedge y \leqslant y' \wedge y' \leqslant y \Rightarrow y \leqslant x$ which together with $x < y \Rightarrow x \leqslant y$ means that $x \sim y$ contradicting $\sim[x] \neq \sim[y]$ so we must have $x' < y'$ proving that the definition is well defined.
3.
   a. If $\sim[x] \in A/\sim$ then $\sim[x] = \sim[x] \Rightarrow \sim[x] \leqslant \sim[x]$
   b. If $\sim[x] \leqslant \sim[y]$ and $\sim[y] \leqslant \sim[x]$ then we have either
      i. $\sim[x] = \sim[y]$ proving anti-symmetry
      ii. $\sim[x] < \sim[y]$ and $\sim[y] < \sim[x] \Rightarrow x < y \wedge y < x \Rightarrow x \leqslant y \wedge y \leqslant x \Rightarrow x \sim y \Rightarrow \sim[x] = \sim[y]$ proving anti-symmetry
   c. If $\sim[x] \leqslant \sim[y] \wedge \sim[y] \leqslant \sim[z]$ then we have the following cases to consider
      i. $\sim[x] = \sim[y] \wedge \sim[y] = \sim[z] \Rightarrow \sim[x] = \sim[z] \Rightarrow \sim[x] \leqslant \sim[z]$
      ii. $\sim[x] = \sim[y] \wedge \sim[y] < \sim[z] \Rightarrow \sim[x] < \sim[z] = \sim[x] \leqslant \sim[z]$
      iii. $\sim[x] < \sim[y] \wedge \sim[y] = \sim[z] \Rightarrow \sim[x] < \sim[z] \Rightarrow \sim[x] \leqslant \sim[z]$
      iv. $\sim[x] < \sim[y] \wedge \sim[y] < \sim[z] \Rightarrow x < y \wedge y < z \Rightarrow x < z \Rightarrow \sim[x] < \sim[y] \Rightarrow \sim[x] \leqslant \sim[y]$
4. If $x \leqslant y \Rightarrow x < y \vee x = y \Rightarrow \sim[x] < \sim[y] \vee \sim[x] = \sim[y] \Rightarrow \sim[x] \leqslant \sim[y]$
5. If $x' \in \sim[x], y' \in \sim[y]$ then $x' \leqslant x \wedge x \leqslant x'$ and $y' \leqslant y \wedge y \leqslant y'$. If $\sim[x] \leqslant \sim[y] \Rightarrow \sim[x] < \sim[y] \vee \sim[x] = \sim[y]$ giving the following possibilities
   a. $x < y \Rightarrow x' \leqslant x < y \leqslant y' \Rightarrow x' \leqslant y'$
   b. $x \leqslant y \wedge y \leqslant x \Rightarrow x' \leqslant x \leqslant y \leqslant y' \Rightarrow x' \leqslant y'$ □

Theorem 2.139. *If $\langle A, \leqslant\rangle$ is a partial ordered (pre-ordered) class and $B \subseteq A$ then we can define the relation $\leqslant_{|B} = \leqslant \bigcap (B \times B) \subseteq B \times B$ we have then that $\langle B, \leqslant_{|B}\rangle$ is a partial ordered (pre-ordered) class. In general we use the simpler notation $\langle B, \leqslant\rangle$ for the order relation (pre-order) on B.*

Proof. The proof is quite trivial

1. $x \in B \Rightarrow x \in A \Rightarrow x \leqslant x \Rightarrow (x, x) \in \leqslant \bigcap (B \times B) = \leqslant_{|B}$
2. $x \leqslant_{|B} y \wedge y \leqslant_{|B} x \Rightarrow x \leqslant y \wedge y \leqslant x \Rightarrow x = y$
3. $x \leqslant_{|B} y \wedge y \leqslant_{|B} z \Rightarrow x, y, z \in B \wedge x \leqslant y \wedge y \leqslant z \Rightarrow x, z \in B \wedge x \leqslant z \Rightarrow x \leqslant_{|B} z$ □

Definition 2.140. *If $\langle A, \leqslant_A\rangle$ and $\langle B, \leqslant_B\rangle$ are partially ordered sets then the lexical order $\leqslant_{A,B} \subseteq (A \times B) \times (A \times B)$ is defined by $\leqslant_{A,B} = \{((x, y), (x', y')) \in (A \times B) \times (A \times B) | ((x \neq x') \wedge (x \leqslant_A x')) \vee ((x = x') \wedge (y \leqslant_B y')) | \}$*

Theorem 2.141. *If $\langle A, \leqslant_A\rangle$ and $\langle B, \leqslant_B\rangle$ are partially ordered sets then $\langle A \times B, \leqslant_{A,B}\rangle$ is a partially ordered set*

Proof.

1. If $(x, y) \in A \times B$ then $x = x \wedge y \leqslant_B y \Rightarrow (x, y) \leqslant_{A,B} (x, y)$

2. If $(x, y) \leqslant_{A,B} (x', y')$ and $(x', y') \leqslant_{A,B} (x, y)$ then if $x \neq x'$ we must have $x \leqslant_A x'$ and $x' \leqslant_A x$ giving $x = x'$ a contradiction. So we must have $x = x'$ but then we have $y \leqslant_B y'$ and $y' \leqslant_B y$ and thus $y = y'$

3. If $(x, y) \leqslant_{A,B} (x', y')$ and $(x', y') \leqslant_{A,B} (x'', y'')$ then consider the following possible cases for x, x'

   a. $x = x' \Rightarrow y \leqslant_B y'$ now we have the following possibilities for x' and x''

      i. $x' = x'' \Rightarrow y' \leqslant_B y'' \Rightarrow x = x'' \wedge y \leqslant_B y'' \Rightarrow (x, y) \leqslant_{A,B} (x'', y'')$

      ii. $x' \neq x'' \Rightarrow x' \leqslant_A x'' \underset{x = x' \Rightarrow x \leqslant_A x'}{\Rightarrow} x \neq x'' \wedge x \leqslant_A x'' \Rightarrow (x, y) \leqslant_{A,B} (x'', y'')$

   b. $x \neq x' \Rightarrow x \leqslant_A x'$ now we have the following possibilities for x' and x''

      i. $x' = x'' \Rightarrow x' \leqslant_A x'' \wedge x \neq x'' \Rightarrow x \neq x'' \wedge x \leqslant_A x'' \Rightarrow (x, y) \leqslant_{A,B} (x'', y'')$

      ii. $x' \neq x'' \Rightarrow x' \leqslant_A x'' \Rightarrow x \leqslant_A x''$ now we have the following cases

         A. $x = x'' \Rightarrow x' \leqslant_A x \Rightarrow x = x'$ a contradiction so this case does not occur

         B. $x \neq x'' \Rightarrow (x, y) \leqslant_{A,B} (x'', y'')$ □

Definition 2.142. *If $\langle A, \leqslant \rangle$ is a partially ordered class then $x, y \in A$ are comparable if $x \leqslant y$ or $y \leqslant x$*

Theorem 2.143. *If $\langle A, \leqslant \rangle$ is a partially ordered class and $x, y \in A$ are comparable then we have either $x \leqslant y$ or $y < x$*

Proof. If x, y are comparable then we have $(x \leqslant y) \vee (y \leqslant x)$ now we have the following exclusive cases

1. $x \leqslant y \Rightarrow x \leqslant y$

2. $\neg(x \leqslant y) \Rightarrow y \leqslant x$ now if $x = y \Rightarrow x \leqslant y$ contradicting $\neg(x \leqslant y)$ so we have $x \neq y$ and this together with $y \leqslant x$ gives $y < x$ □

Definition 2.144. *A partially ordered (pre-ordered) class $\langle A, \leqslant \rangle$ is called a* ***fully ordered class (fully pre-ordered)*** *or a* ***linear ordered class (linear pre-ordered class)*** *if $\forall x, y \in A \vDash x \leqslant y \vee y \leqslant x$. A subclass $B \subseteq A$ is called a* ***fully ordered (pre-ordered) subclass of*** *A or a* ***linear ordered (pre-ordered) subclass*** *or a* ***chain*** *of A if $\forall x, y \in B \vDash x \leqslant y \vee y \leqslant x$. So a fully ordered (sub)class is a (sub)class where every element is comparable with every other element. Also in a fully ordered class we have for every x, y either $x \leqslant y$ or $y < x$ (see previous theorem).*

Theorem 2.145. *Let $\langle A, \leqslant \rangle$ be a totally ordered class, $B \subseteq A$ then $\langle B, \leqslant_{|B} \rangle$ is totally ordered (see 2.139)*

Proof. Using 2.139 we have that $\langle B, \leqslant_{|B} \rangle$ is partially ordered. If now $x, y \in B \subseteq A$ then by the total ordering of $\leqslant$ we have that $(x, y) \in \leqslant$ or $(y, x) \in \leqslant$ which as $(x, y), (y, x) \in B \times B$ is equivalent with $(x, y) \in \leqslant \bigcap B \times B$ or $(y, x) \in \leqslant \bigcap B \times B$. □

Theorem 2.146. *Let $\langle A, \leqslant_A \rangle$ and $\langle B, \leqslant_B \rangle$ be fully ordered classes then $\langle A \times B, \leqslant_{A,B} \rangle$ is a fully ordered class (see 2.141)*

Proof. First $\langle A \times B, \leqslant_{A,B} \rangle$ is a partially ordered class by 2.141. If now (x, y), $(x', y') \in A \times B$ then we have for x, x' either

- $\boldsymbol{x = x'}$**.** As $\langle B, \leqslant_B \rangle$ is fully ordered we have either
  - $\boldsymbol{y \leqslant y'}$**.** then $(x, y) \leqslant (x', y')$
  - $\boldsymbol{y' < y}$**.** then $(x', y') \leqslant (x, y)$
- $\boldsymbol{x \neq x'}$**.** As $\langle A, \leqslant_A \rangle$ is fully ordered we have either
  - $\boldsymbol{x \leqslant x'}$**.** then $(x, y) \leqslant (x', y')$
  - $\boldsymbol{x' \leqslant x}$**.** then $(x', y') \leqslant (x, y)$ □

Definition 2.147. *If $\langle A, \leqslant \rangle$ is a partially ordered class and $a \in A$ then the* ***initial segment of A determined by a*** *S_a is defined by*

$$S_a = \{x \in A | x < a\}$$

Theorem 2.148. *If $\langle A. \leqslant \rangle$ and P is a initial segment of A and Q a initial segment of P (using the induced order relation on P) then Q is a initial segment of A*

Proof. By the hypothesis of the theorem we have $\exists a \in A \vdash P = \{x \in A | x < a\}$ and $\exists b \in P \vdash Q = \{x \in P \vdash x <_{|P} b\}$ then we have as $b \in P \Rightarrow b < a$

$$\begin{array}{rcl}
x \in S_b = \{x \in A | x < b\} & \Leftrightarrow & x \in A \wedge x < b \\
& \underset{x<b<a \Rightarrow x<a}{\Rightarrow} & x \in A \wedge x < a \wedge x < b \\
& \Rightarrow & x \in P \wedge x < b \\
& \Rightarrow & x \in P \wedge x <_{|P} b \\
& \Rightarrow & x \in Q \\
x \in Q & \Rightarrow & x \in P \wedge x <_{|p} b \\
& \Rightarrow & x \in P \wedge x < b \\
& \underset{P \subseteq A}{\Rightarrow} & x \in A \wedge x < b \\
& \Rightarrow & x \in S_b
\end{array}$$

□

Definition 2.149. *If $\langle A, \leqslant \rangle$ is a partially ordered class then a* ***cut*** *of A is a pair $\langle C, D \rangle$ such that*

1. $C \neq \emptyset \wedge D \neq \emptyset$

2. $C \bigcap D = \emptyset$
3. $x \in C \wedge y \leqslant x \Rightarrow y \in C$
4. $x \in D \wedge y \geqslant x \Rightarrow y \in D$

Definition 2.150. *If $\langle A, \leqslant_A \rangle$ and $\langle B, \leqslant_B \rangle$ are partially ordered classes then a function $f: A \rightarrow B$ is **increasing** if $\forall x, y \in A$ we have*

$$x \leqslant y \ \Rightarrow \ f(x) \leqslant f(y)$$

*Such a function is also called **order preserving**.*

Theorem 2.151. *If $\langle A, \leqslant_A \rangle$, $\langle B, \leqslant_B \rangle$ and $\langle C, \leqslant_C \rangle$ are partially ordered classes then $f: A \rightarrow B$ and $g: B \rightarrow C$ are increasing functions then $g \circ f$ is a increasing function. Further if f and g is strictly increasing then $g \circ f$ is strictly increasing.*

Proof. If $x \leqslant y \Rightarrow f(x) \leqslant f(y) \Rightarrow g(f(x)) \leqslant g(f(y)) \Rightarrow (g \circ f)(x) \leqslant (g \circ f)(y)$. If g is strictly increasing then $x < y \Rightarrow f(x) < f(y) \Rightarrow g(f(x)) < g(f(y)) \Rightarrow (g \circ f)(x) < (g \circ f)(y)$ □

Definition 2.152. *If $\langle A, \leqslant_A \rangle$ and $\langle B, \leqslant_B \rangle$ are partially ordered classes then a function $f: A \rightarrow B$ is **strictly increasing** if $\forall x, y \in A$ we have*

$$x < y \ \Rightarrow \ f(x) < f(y)$$

Definition 2.153. *If $\langle A, \leqslant_A \rangle$ and $\langle B, \leqslant_B \rangle$ are partially ordered classes then a function $f: A \rightarrow B$ is a **isomorphism** iff f is bijective and $\forall x, y \in A$ we have*

$$x \leqslant_A y \ \Leftrightarrow \ f(x) \leqslant_B f(y)$$

Notation 2.154. *If $\langle A, \leqslant_A \rangle$ and $\langle B, \leqslant_B \rangle$ are partially ordered class and there exists a isomorphism $f: A \rightarrow B$ between A and B then we say that A and B are isomorphic noted by $A \cong B$*

Theorem 2.155. *If $\langle A, \leqslant_A \rangle$ and $\langle B, \leqslant_B \rangle$ are partially ordered classes and f: $A \rightarrow B$ a **isomorphism** then*

$$x <_A y \ \Leftrightarrow \ f(x) <_B f(y)$$

Proof.

$\Rightarrow$

If $x <_A y \Rightarrow x \leqslant_A y \wedge x \neq y \Rightarrow f(x) \leqslant_B f(y)$. If $f(x) = f(y) \underset{f \text{ is injective}}{\Rightarrow} x = y$ contradicting $x \neq y$ so we have $f(x) \neq f(y)$ and thus $f(x) <_B f(y)$

$\Leftarrow$

If $f(x) <_B f(y) \Rightarrow f(x) \leqslant_B f(y) \underset{f \text{ is isomorphism}}{\Rightarrow} x \leqslant_A y$ now if $x = y \Rightarrow f(x) = f(y)$ contradicting $f(x) <_B f(y) \Rightarrow x <_A y$ □

Theorem 2.156. *If $\langle A, \leqslant_A \rangle$ and $\langle B, \leqslant_B \rangle$ are partially ordered classes then a bijective function $f: A \rightarrow B$ is a isomorphism iff $f: A \rightarrow B$ and $f^{-1}: B \rightarrow A$ are increasing functions*

Proof. First as f is bijective f^{-1} exists and

1. If f is a isomorphism then $x \leqslant_A y \Rightarrow f(x) \leqslant_B f(y) \Rightarrow f$ is a increasing function. If $x, y \in B$ with $x \leqslant_A y$ then $f(f^{-1}(x)) = x \leqslant_A y = f(f^{-1}(y)) \underset{f \text{ is a isomorphism}}{\Rightarrow} f^{-1}(x) \leqslant_B f^{-1}(y)$ so f^{-1} is a increasing function.

2. Suppose f, f^{-1} are increasing functions then if $x \leqslant_A y \underset{f \text{ is increasing}}{\Rightarrow} f(x) \leqslant_B f(y)$. Suppose that $f(x) \leqslant_B f(y) \underset{f^{-1} \text{ is increasing}}{\Rightarrow} f^{-1}(f(x)) \leqslant_A f^{-1}(f(y)) \Rightarrow x \leqslant y$ □

Theorem 2.157. *If $\langle A, \leqslant_A\rangle$, $\langle B, \leqslant_B\rangle$ and $\langle C, \leqslant_C\rangle$ are partially ordered classes then*

1. *$1_A: A \to A$ is a isomorphism*
2. *If $f: A \to B$ is a isomorphism then $f^{-1}: B \to A$ is a isomorphism*
3. *If $f: A \to B$ and $g: B \to C$ are isomorphism's then $g \circ f$ is a isomorphism*

Proof.

1. By 2.24 we have that $1_A: A \to A$ is a bijection then as $x = 1_A(x)$ and $y = 1_A(y)$ so that $x \leqslant y \Leftrightarrow 1_A(x) \leqslant 1_A(y)$.

2. If $f: A \to B$ is a isomorphism then by 2.38 we have that $f^{-1}: B \to A$ is a bijection. By the previous theorem we have that f^{-1} is increasing. So assume $f^{-1}(x) \leqslant f^{-1}(y) \underset{f \text{ is increasing}}{\Rightarrow} f(f^{-1}(x)) \leqslant f(f^{-1}(y)) \Rightarrow x \leqslant y$.

3. Using 2.46 we have that $g \circ f$ is a bijection and that $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. Now using 2.151 and the fact that f, g, f^{-1}, g^{-1} are increasing we have that $g \circ f$ and $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$ are increasing. Using 2.156 we have then that $g \circ f$ is a isomorphism. □

Theorem 2.158. *If $\langle A, \leqslant_A\rangle$, $\langle B, \leqslant_B\rangle$ and $\langle C, \leqslant_C\rangle$ are partially ordered classes then we have*

1. *$A \cong A$*
2. *If $A \cong B$ then $B \cong A$*
3. *If $A \cong B$ and $B \cong D$ then $B \cong D$*

Proof. This follows easily from the previous theorem. □

Theorem 2.159. *If $\langle A, \leqslant_A\rangle$ is a fully ordered class and $\langle B, \leqslant_B\rangle$ is a partially ordered class then a bijective and increasing function $f: A \to B$ is a isomorphism*

Proof. Suppose that $f(x) \leqslant_B f(y)$ then since A is fully ordered we have that x, y are comparable therefore by 2.143 we have the following exclusive cases

1. $x \leqslant_A y$ in this case or theorem is proved

2. $y <_A x$ in this case we would have $f(y) \leqslant_B f(x) \Rightarrow f(y) = f(x) \underset{f \text{ is injective}}{\Rightarrow} x = y$ a contradiction. So this case does not occurs. □

Definition 2.160. *If $\langle X, \leqslant\rangle$ is a pre-ordered class and $A \leqslant X$ then*

1. *m is a maximal element of A iff $m \in A$ and if $x \in A$ with $m \leqslant x$ then $x \leqslant m$*
2. *m is a minimal element of A iff $m \in A$ and if $x \in A$ with $x \leqslant m$ then $m \leqslant x$*

Definition 2.161. *If $\langle X, \leqslant\rangle$ is a partially ordered class and $A \subseteq X$ then*

1. *m is a maximal element of A iff $m \in A$ and $\forall x \in A$ if $x \geqslant m \Rightarrow x = m$*
2. *m is a minimal element of A iff $m \in A$ and $\forall x \in A$ if $x \leqslant m \Rightarrow x = m$*
3. *m is the greatest element of A iff $m \in A$ and $\forall x \in A \vdash x \leqslant m$*
4. *m is the least element of A iff $m \in A$ and $\forall x \in A \vdash m \leqslant x$*

note that (3) and (4) suggests that there is only one greatest and least element, this is indeed true as is proved in the next theorem.

Note 2.162. There is a difference between the definition of a maximal element and a greatest element. If m is a greatest element of A then m is comparable with every element in A. This is not needed if m is a maximal element of A. The same goes for minimal element and least element.

Theorem 2.163. *If $\langle X, \leqslant\rangle$ is a partially ordered class and $A \subseteq X$ then if A has a greatest element then this is unique. If A has a least element then it is unique. The unique greatest element of A if it exists is called the maximum of A and noted by* $\max(A)$*. The unique least element if it exists is called the minimum of A and noted by* $\min(A)$*.*

Note: The existence of a maximum is not the same as the existence of a maximal element, the existence of a minimum is not the same as the existence of a minimal element.

Proof.

1. If m, m' are greatest elements then $m, m' \in A \Rightarrow m \leqslant m' \wedge m' \leqslant m \Rightarrow m = m'$
2. If m, m' are least elements then $m, m' \in A \Rightarrow m \leqslant m' \wedge m' \leqslant m \Rightarrow m = m'$ □

Theorem 2.164. *Let $\langle X, \leqslant\rangle$ be a partially ordered class and let $B, C \subseteq X$ sub-classes such that* $\max(B), \max(C)$ *exists [or* $\min(B)\min(C)$ *exists] then if $\forall x \in B$ there exists a $y \in C$ with $x \leqslant y$ [or $\forall x \in B$ there exists a $y \in C$ such that $y \leqslant x$] then* $\max(B) \leqslant \max(C)$ *[or* $\min(C) \leqslant \min(B)$*]*

Proof. If $\max(B), \max(C)$ exists then $\max(B) \in B$ and thus there exist a $c \in C$ such that $\max(B) \leqslant c \leqslant \max(C) \Rightarrow \max(B) \leqslant \max(C)$. If $\min(B), \min(C)$ exists then as $\min(B) \in B$ there exists a $c \in C$ such that $\min(C) \leqslant c \leqslant \min(B) = \min(C) \leqslant \min(B)$. □

Definition 2.165. *If $\langle A, \leqslant\rangle$ is a pre-ordered class and $B \subseteq A$ then*

1. *$u \in A$ is a upper bound of B iff $\forall x \in B \vdash x \leqslant u$*
2. *B is bounded above if it has a upper bound*
3. *$l \in A$ is a lower bound of B iff $\forall x \in B \vdash l \leqslant x$*

4. *B is bounded below if it has a lower bound*

Definition 2.166. *If $\langle A, \leqslant \rangle$ is a partially ordered class and $B \subseteq A$ a subclass of A then*

1. *$u \in A$ is a upper bound of B iff $\forall x \in B \vdash x \leqslant u$*
2. *B is bounded above if it has a upper bound*
3. *$l \in A$ is a lower bound of B iff $\forall x \in B \vdash l \leqslant x$*
4. *B is bounded below if it has a lower bound*
5. *$\upsilon(B) = \{u \in A | u \text{ is a upper bound of } B\}$ (the class of all the upper bounds of B)*
6. *$\lambda(B) = \{l \in A | l \text{ is a lower bound of } B\}$ (the class of all the lower bounds of B)*
7. *If $\upsilon(B)$ has a least element then this is called the **least upper bound** or **supremum** of B which is noted as $\sup(B)$ (it is unique by the previous theorem).*
8. *If $\lambda(B)$ has a greatest element then this is called the **greatest lower bound** or **infinum** of B which is noted as $\inf(B)$ (it is unique by the previous theorem).*

Theorem 2.167. *If $\langle A, \leqslant \rangle$ is a partially ordered class and $B \subseteq A$ then we have*

1. *If $\inf(B)$ exists then $\forall a \in A$ with $\inf(B) < a$ there exists a $b \in B$ such that $\inf(B) \leqslant b < a$*
2. *If $\sup(B)$ exists then $\forall a \in A$ with $a < \sup(B)$ there exists a $b \in B$ such that $a < b \leqslant \sup(B)$*

Proof.

1. We proceed by contradiction so assume that $\exists a \in A$ with $\inf(B) < a$ so that $\forall b \in B$ we have that $\neg(\inf(B) \leqslant b < a) = \neg(\inf(B) \leqslant b \wedge b < a) = b < \inf(B) \vee a \leqslant b \underset{\inf(B) \text{ is a lower bound of } B \Rightarrow \inf(B) \leqslant b}{\Rightarrow} a \leqslant b \Rightarrow a$ is a lower bound of B and thus by definition of $\inf(B)$ we have $a \leqslant \inf(B)$ contradicting $\inf(B) < a$. So we must have that $\forall a \in A$ with $\inf(B) < a$ there exist a $b \in B$ such that $\inf(B) \leqslant b < a$.
2. We proceed by contradiction so assume that $\exists a \in A$ with $a < \sup(B)$ so that $\forall b \in B$ we have that $\neg(a < b \leqslant \sup(B)) = \neg(a < b \wedge b \leqslant \sup(B)) = b \leqslant a \vee \sup(B) < b \underset{\sup(B) \text{ is a upper bound of } B \Rightarrow b \leqslant \sup(B)}{\Rightarrow} b \leqslant a \Rightarrow a$ is a upper bound of B and thus by definition of $\sup(B)$ we have $\sup(B) \leqslant a$ contradicting $a < \sup(B)$. So we must have that $\forall a \in A$ with $a < \sup(B)$ there exist a $b \in B$ such that $a < b \leqslant \sup(B)$. □

Theorem 2.168. *Let $\mathcal{A}$ be a class of classes partially ordered by inclusion ($\langle \mathcal{A}, \subseteq \rangle$) then if $\mathcal{B} \subseteq \mathcal{A}$ then*

1. *If $\bigcap_{B \in \mathcal{B}} B \in \mathcal{A} \Rightarrow \inf(\mathcal{B}) = \bigcap_{B \in \mathcal{B}} B$*

2. *If* $\bigcup_{B\in\mathcal{B}} B\in\mathcal{A}\Rightarrow\sup(\mathcal{B})=\bigcup_{B\in\mathcal{B}} B$

Proof.

1. If $B\in\mathcal{B}\Rightarrow\bigcap_{B\in\mathcal{B}} B\subseteq B\Rightarrow\bigcap_{B\in\mathcal{B}} B\in\lambda(\mathcal{B})$ now if $C\in\lambda(\mathcal{B})$ then $\forall B\in\mathcal{B}$ we have $C\subseteq B\Rightarrow C\subseteq\bigcap_{B\in\mathcal{B}} B$ so $\bigcap_{B\in\mathcal{B}} B$ is the greatest element of $\lambda(B)$
2. If $B\in\mathcal{B}\Rightarrow B\subseteq\bigcup_{B\in\mathcal{B}} B\Rightarrow\bigcup_{B\in\mathcal{B}} B\in\upsilon(\mathcal{B})$ now if $C\in\upsilon(\mathcal{B})$ then $\forall B\in\mathcal{B}$ we have $B\subseteq C\Rightarrow\bigcup_{B\in\mathcal{B}} B\subseteq C$ so $\bigcup_{B\in\mathcal{B}} B$ is the least element of $\upsilon(\mathcal{B})$ □

Theorem 2.169. *If* $\langle A,\leqslant\rangle$ *is a partially ordered class then if* $B,C\subseteq A$ *with* $B\subseteq C$ *we have*

1. *If* B *has a greatest element* b *and* C *has a greatest element* c *then* $b\leqslant c$
2. *If* B *has a least element* b *and* C *has a least element* c *then* $c\leqslant b$

Proof.

1. $b\in B\Rightarrow b\in C\underset{\forall x\in C\vdash x\leqslant c}{\Rightarrow} b\leqslant c$
2. $b\in B\Rightarrow b\in C\underset{\forall x\in C\vdash c\leqslant x}{\Rightarrow} c\leqslant b$ □

Theorem 2.170. *If* $\langle A,\leqslant\rangle$ *is a partially ordered class then if* $B,C\subseteq A$ *with* $B\subseteq C$ *we have*

1. $\upsilon(C)\subseteq\upsilon(B)$
2. $\lambda(C)\subseteq\lambda(B)$

Proof.

1. If $x\in\upsilon(C)$ then $\forall b\in B\vdash b\in C\Rightarrow b\leqslant x\Rightarrow x\in\upsilon(B)$
2. If $x\in\lambda(C)$ then $\forall b\in B\vdash b\in C\Rightarrow x\leqslant b\Rightarrow x\in\lambda(B)$ □

Theorem 2.171. *If* $\langle A,\leqslant\rangle$ *is a partially ordered class then if* $B,C\subseteq A$ *with* $B\subseteq C$

1. If the supremum exists for B,C then $\sup(B)\leqslant\sup(C)$
2. If the infinum exists for B,C then $\inf(C)\leqslant\inf(B)$

Proof.

1. By 2.170 we have $\upsilon(C)\subseteq\upsilon(B)$ so using 2.169 we have $\sup(B)$ = least element of $\upsilon(B)\leqslant$ least element of $\upsilon(C)=\sup(C)$
2. By 2.170 we have $\lambda(C)\subseteq\lambda(B)$ so using 2.169 we have $\inf(C)=$ greatest element of $\lambda(C)\leqslant$ greatest element of $B=\inf(B)$ □

Theorem 2.172. *If* $\langle A,\leqslant\rangle$ *is a partially ordered class,* $B,C\subseteq A$ *are subsets of* A *such that* $\forall b\in B$ *there exists a* $c\in C$ *such that* $b\leqslant c$ *then* $\sup(B)\leqslant\sup(C)$ *(assuming that the supremum of* B,C *exists). On the other hand if* $\forall b\in B$ *there exists a* $c\in C$ *such that* $c\leqslant b$ *then* $\inf(C)\leqslant\inf(B)$ *(assuming that the infinum of* B,C *exits).*

Proof. If $b \in B$ then there exists a $c \in C$ such that $b \leqslant c \leqslant \sup(C)$ ($\sup(C)$ is a upper bound of C) so that $\sup(C)$ is a upper bound of B and thus we have $\sup(B) \leqslant \sup(C)$.

On the other hand if $b \in B$ then there exists a $c \in C$ such that $\inf(C) \leqslant c \leqslant b$ so that $\inf(C)$ is a lower bound of B so that $\inf(C) \leqslant \inf(B)$ □

Theorem 2.173. *If $\langle A, \leqslant\rangle$ is a partially ordered class with $B \subseteq A$ then $B \subseteq \upsilon(\lambda(B))$ and $B \subseteq \lambda(\upsilon(B))$*

Proof.

1. If $x \in B \Rightarrow \forall y \in \lambda(B) \vdash y \leqslant x \Rightarrow x \in \upsilon(\lambda(B))$
2. If $x \in B \Rightarrow \forall y \in \upsilon(B) \vdash x \leqslant y \Rightarrow x \in \lambda(\upsilon(B))$ □

Theorem 2.174. *If $\langle A, \leqslant\rangle$ is a partially ordered class with $B \subseteq A$ a subclass then*

1. *If $\lambda(B)$ has a supremum then B has a infinum and $\sup(\lambda(B)) = \inf(B)$*
2. *If $\upsilon(B)$ has a infinum then B has a supremum and $\inf(\upsilon(B)) = \sup(B)$*

Proof.

1. Let $s = \sup(\lambda(B))$ then if $b \in B$ we have $\forall y \in \lambda(B)$ that $y \leqslant b$ hence $b \in \upsilon(\lambda(B))$ thus (s is least upper bound of $\lambda(B)$) $s \leqslant b$. As b has been chosen arbitrary we have that $s \in \lambda(B)$. Now if $d \in \lambda(B)$ then $d \leqslant s$ [because $s = \sup(\lambda(B)) \Rightarrow s$ is the least element of $\upsilon(\lambda(B)) \Rightarrow s \in \upsilon(\lambda(B))$] so s is the greatest element of $\lambda(B)$ hence $s = \inf(B)$.
2. Let $s = \inf(\upsilon(B))$ then if $b \in B$ we have $\forall y \in \upsilon(B)$ that $b \leqslant y$ hence $b \in \lambda(\upsilon(B))$ thus (s is greatest lower bound of $\upsilon(B)$) $b \leqslant s$. As b has been chosen arbitrary we have $s \in \upsilon(B)$. Now if $d \in \upsilon(B)$ then $s \leqslant d$ [because $s = \inf(\upsilon(B)) \Rightarrow s$ is the greatest element of $\lambda(\upsilon(B)) \Rightarrow s \in \lambda(\upsilon(B))$] so s is the least element of $\upsilon(B)$ hence $s = \sup(B)$ □

Definition 2.175. *If $\langle A, \leqslant\rangle$ is a partially ordered class. If every nonempty subclass of A that is bounded above has a supremum then A is said to be **conditional complete**.*

Next theorem shows that conditional completeness can be defined also with lower bounds and infinum

Theorem 2.176. *If $\langle A, \leqslant\rangle$ is a partially ordered class then the following are equivalent*

1. *Every nonempty subclass of A that is bounded above has a supremum*
2. *Every nonempty subclass of A that is bounded below has a infinum*

Proof.

$1 \Rightarrow 2$

Let $B \subseteq A$ which is bounded below$\Rightarrow \lambda(B) \neq \emptyset$. Now as $B \neq \emptyset$ there exists a $b \in B$ and by definition of $\lambda(B)$ we have $\forall y \in \lambda(B)$ that $y \leqslant b$. So $\lambda(B)$ is nonempty and bounded above. From the hypothesis we have then that $\sup(\lambda(B))$ exists but then by 2.174 we have that B has a infinum.

$2 \Rightarrow 1$

Let $B \subseteq A$ which is bounded above$\Rightarrow \upsilon(B) \neq \emptyset$. Now as $B \neq \emptyset$ there exists a $b \in B$ and by definition of $\upsilon(B)$ we have $\forall y \in \upsilon(B)$ that $b \leqslant y$. So $\upsilon(B)$ is nonempty and bounded below. From the hypothesis we have then that in$(\upsilon(B))$ exists and then by 2.174 we have that B has a supremum. □

We show now that a order preserving isomorphism conserves the supremum and infinum

Lemma 2.177. *Let $\langle A, \leqslant_A \rangle$, $\langle B, \leqslant_B \rangle$ be partially ordered classes and $f: A \rightarrow B$ a order preserving isomorphism (see isomorphism) then*

1. *If b is a upper bound of $f(S)$ then $(f^{-1})(b)$ is a upper bound of $f^{-1}(S)$*
2. *If l is a lower bound of $f(S)$ then $(f^{-1})(l)$ is a lower bound of $f^{-1}(S)$*
3. *If $S \subseteq A$ so that $\sup(S)$ exists then $\sup(f(S))$ exists and is equal to $f(\sup(\mathcal{S}))$*
4. *If $S \subseteq A$ so that $\inf(S)$ exists then $\inf(f(S))$ exists and is equal to $f(\inf(S))$*

Proof. As f is a isomorphism we have using 2.156 that f and f^{-1} are increasing functions.

1. Let b be a upper bound of $f(S)$ then if $a \in f^{-1}(S)$ we have $f(a) \in S$ so that $f(a) \leqslant_B b$ and thus $a = f^{-1}(f(a)) \leqslant_A f^{-1}(b)$ proving that $f^{-1}(b)$ is a upper bound
2. Let l be a lower bound of $f(S)$ then if $a \in f^{-1}(S)$ we have $f(a) \in S$ so that $l \leqslant_B f(a)$ and thus $f^{-1}(l) \leqslant_A f^{-1}(f(a)) = a$ proving that $f^{-1}(l)$ is a lower bound.
3. If $S \subseteq A$ so that $\sup(S)$ exist then if $a \in f(S)$ there exists $a' \in S$ such that $a = f(a')$, as $\sup(S)$ is a upper bound of S we have that $a' \leqslant_A \sup(S)$ so that $a = f(a') \leqslant_B f(\sup(S))$ proving that
$$f(\sup(S)) \text{ is a upper bound of } f(S) \tag{2.18}$$
If now b is a upper bound of $f(S)$ then using (1) $(f^{-1})(b)$ is a upper bound of $f^{-1}(f(S)) \underset{f \text{ is a bijection}}{=} S$ hence $\sup(S) \leqslant_A (f^{-1})(b)$ so that $f(\sup(S)) \leqslant_B f((f^{-1})(b)) = b$ which proves together with 2.18 that $f(\sup(S)) = \sup(f(S))$
4. If $S \subseteq A$ so that $\inf(S)$ exists then if $a \in f(S)$ there exists a $a' \in \mathcal{S}$ such that $a = f(a')$, as $\inf(S)$ is a lower bound of S we have that $\inf(S) \leqslant_A a'$ so that $f(\sup(S)) \leqslant_B f(a') = a$ proving that
$$f(\inf(S)) \text{ is a lower bound of } f(S) \tag{2.19}$$
If now l is a lower bound of $f(S)$ then using (2) $(f^{-1})(b)$ is a lower bound of $f^{-1}(f(S)) \underset{f \text{ is a bijection}}{=} S$ hence $(f^{-1})(b) \leqslant_A \inf(S)$ so that $b = f((f^{-1})(b)) \leqslant_B f(\inf(S))b$ which proves together with 2.18 that $f(\inf(S)) = \inf(f(S))$ □

Corollary 2.178. *Let $\langle A, \leqslant_A \rangle$ be a conditionally complete partial ordered set, $\langle B, \leqslant_B \rangle$ a partially ordered set and $f: A \rightarrow B$ a isomorphism (see isomorphism) then $\langle B, \leqslant_B \rangle$ is conditionally complete.*

Proof. Let $\emptyset \neq S \subseteq B$ be a non empty set such that there exists a upper bound b then using the previous theorem (see 2.177) we have that $f^{-1}(b)$ is a upper bound of $f^{-1}(S)$. As $S \neq \emptyset$ there exists a $a \in S$ and as f is a bijection there exists a $a' \in A$ such that $f(a') = a$ hence $a' \in f^{-1}(S)$. From the conditionally completeness of $\langle A, \leqslant_A \rangle$ we have that $\sup(f^{-1}(S))$ exists. Using 2.177 we have that $S = f(f^{-1}(S))$ has a supremum. □

Definition 2.179. *Let $\langle A, \leqslant \rangle$ be a pre-ordered set then if $a, b \in A$ with $a \leqslant b$ we define*

1. $[a, b] = \{x \in A | a \leqslant x \wedge x \leqslant b\}$
2. $]a, b] = \{x \in A | a < x \wedge x \leqslant b\}$
3. $[a, b[= \{x \in A | a \leqslant x \wedge x < b\}$
4. $]a, b[= \{x \in A | a < x \wedge x < b\}$
5. $]-\infty, b] = \{x \in A | x \leqslant b\}$
6. $]-\infty, b[= \{x \in A | x < b\}$
7. $[a, \infty[= \{x \in A | a \leqslant x\}$
8. $]a, \infty[= \{x \in A | a < x\}$
9. $]-\infty, \infty[= A$
10. $]a, a[= \emptyset$
11. $[a, a] = \{a\}$

*If a set takes one of the forms (1)-(11) it is called a **generalized interval**.*

We have a much simpler definition for a generalized interval if $X$is conditional complete.

Theorem 2.180. *Let $\langle A, \leqslant \rangle$ be a partial ordered class that is conditional complete then the following are equivalent for a subset $B \subseteq A$*

1. *B is a generalized interval*
2. *$\forall x, y \in B$ with $x < y$ we have $[x, y] \subseteq B$*
3. *$\forall x, y \in B$ with $x < y$ we have that $\forall z \in A$ with $x \leqslant z \leqslant y$ that $z \in B$*
4. *$\forall x, y \in B$ with $x < y$ we have that $\forall z \in A$ with $x < z < y$ that $z \in B$*

Proof. The proof is easy but rather tedious

1. **$(1 \Rightarrow 2)$** Let B be a generalized interval then we have either
    a. **$(B = [a, b])$** if $x, y \in [a, b]$ then $a \leqslant x, y \leqslant b$ so if $t \in [x, y] \Rightarrow x \leqslant t \wedge t \leqslant y$ so that $a \leqslant t \wedge t \leqslant b \Rightarrow t \in [a, b] = B \Rightarrow [x, y] \subseteq B$

b. $(\boldsymbol{B}=[\boldsymbol{a},\boldsymbol{b}[)$ if $x,y\in[a,b[$ then $a\leqslant x,y<b$ so if $t\in[x,y]\Rightarrow x\leqslant t\wedge t\leqslant y$ so that $a\leqslant t\wedge t<b\Rightarrow t\in[a,b[=B\Rightarrow[x,y]\subseteq B$

c. $(\boldsymbol{B}=]\boldsymbol{a},\boldsymbol{b}])$ if $x,y\in]a,b]$ then $a<x,y\leqslant b$ so if $t\in[x,y]\Rightarrow x\leqslant t\wedge t\leqslant y$ so that $a<t\wedge t\leqslant b\Rightarrow t\in]a,b]=B\Rightarrow[x,y]\subseteq B$

d. $(\boldsymbol{B}=]\boldsymbol{a},\boldsymbol{b}[)$ if $x,y\in]a,b[$ then $a<x,y<b$ so if $t\in[x,y]\Rightarrow x\leqslant t\wedge t\leqslant y$ so that $a<t\wedge t<b\Rightarrow t\in]a,b[=B\Rightarrow[x,y]\subseteq B$

e. $(\boldsymbol{B}=]-\boldsymbol{\infty},\boldsymbol{a}])$ if $x,y\in]-\infty,a]$ then $x,y\leqslant a$ so if $t\in[x,y]\Rightarrow x\leqslant t\wedge t\leqslant y\Rightarrow t\leqslant a\Rightarrow t\in]-\infty,a]=B\Rightarrow[x,y]\subseteq B$

f. $(\boldsymbol{B}=]-\boldsymbol{\infty},\boldsymbol{a}[)$ if $x,y\in]-\infty,a[$ then $x,y<a$ so if $t\in[x,y]\Rightarrow x\leqslant t\wedge t\leqslant y\Rightarrow t<a\Rightarrow t\in]-\infty,a[=B\Rightarrow[x,y]\subseteq B$

g. $(\boldsymbol{B}=[\boldsymbol{a},\boldsymbol{\infty}[)$ if $x,y\in[a,\infty[$ then $a\leqslant x,y$ so if $t\in[x,y]\Rightarrow x\leqslant t\wedge t\leqslant y\Rightarrow a\leqslant t\Rightarrow t\in[a,\infty[=B\Rightarrow[x,y]\subseteq B$

h. $(\boldsymbol{B}=]\boldsymbol{a},\boldsymbol{\infty}[)$ if $x,y\in]a,\infty[$ then $a<x,y$ so if $t\in[x,y]\Rightarrow x\leqslant t\wedge t\leqslant y\Rightarrow a<t\Rightarrow t\in]a,\infty[=B\Rightarrow[x,y]\subseteq B$

i. $(\boldsymbol{B}=]-\boldsymbol{\infty},\boldsymbol{\infty}[)$ this is trivial as $[x,y]=\{t\in A|x\leqslant t\wedge t\leqslant y\}$

j. $(\boldsymbol{B}=\emptyset)$ then (2) is satisfied vacuously

k. $(\boldsymbol{B}=\{\boldsymbol{a}\})$ then if $x,y\in B$ then $x=y$ so there is not $x<y$ and again (2) is satisfied vacuously

2. **(2 ⇒ 3)** If $x,y\in B$ with $x<y$ then by (2) we have $[x,y]\subseteq B\Rightarrow\forall z\in B$ such that $x\leqslant z\wedge z\leqslant y$ we have $z\in[x,y]\Rightarrow z\in B$

3. **(3 ⇒ 4)** If $x,y\in B$ with $x<y$ and $x<z<y\Rightarrow x\leqslant z\leqslant y\underset{(3)}{\Rightarrow}z\in B$

4. **(4 ⇒ 1)** Consider the following cases for B

   a. $(\boldsymbol{B}=\emptyset)$ then B is a generalized interval

   b. $(\boldsymbol{B}=\{\boldsymbol{b}\})$ then B is a generalized interval

   c. **(B contains at least two elements)** then there exists b_1,b_2 with $b_1<b_2$ then we have the following possibilities for B:

      i. **(B has no upper/lower bound)** Then if $x\in A$ we have the following possibilities related to a_1,a_2

         A. $(\boldsymbol{x}\in[\boldsymbol{a_1},\boldsymbol{a_2}])$ then we have either

            1. $(\boldsymbol{x}=\boldsymbol{a_1})\Rightarrow x\in B$

            2. $(\boldsymbol{x}=\boldsymbol{a_2})\Rightarrow x\in B$

            3. $(\boldsymbol{a_1}<\boldsymbol{x}\wedge\boldsymbol{x}<\boldsymbol{a_2})\underset{(4)}{\Rightarrow}x\in B$

         B. $(\boldsymbol{x}<\boldsymbol{a_1})$ then as there is no lower bound for B $\exists b\in B$ such that $b<x$ (otherwise x is a lower-bound of B) so $b<a_1$ and $b<x<a_1\underset{(4)}{\Rightarrow}x\in B$

C. $(\boldsymbol{a_2 < x})$ then as there is no lower bound for B $\exists b \in B$ such that $x < b$ then as $a_2 < b$ and $a_2 < x < b$ we have by (4) that $x \in B$

So in all cases we have $x \in B$ proving that $A \subseteq B$ and as $B \subseteq A$ we have $B = A$ a generalized interval.

ii. **($\boldsymbol{B}$ has $\boldsymbol{a}$ upper and lower bound)** As A is conditionally complete there exists then $s = \sup(B)$ and $i = \inf(B)$ and as $i \leqslant b_1$, $b_1 \leqslant s$ we have $i \leqslant s$. Consider now the following possibilities.

A. $(\boldsymbol{i, s \in B})$ If $x \in [i, s]$ then either $x = i \in B$, $x = s \in B$ or $i < x < s \Rightarrow i < s$ and $i < x < s \underset{(4)}{\Rightarrow} x \in B$, proving that $[i, s] \subseteq B$. If $x \in B$ then $i \leqslant x \wedge x \leqslant s$ so that $B \subseteq [i, s]$. This proves that $B = [i, s]$ a generalized interval.

B. $(\boldsymbol{i \notin B \wedge s \in B})$ If $x \in]i, s]$ then either $x = s \in B$ or $i < x < s \Rightarrow i < s \wedge i < x < s \underset{(4)}{\Rightarrow} x \in B$ proving that $]i, s] \subseteq B$. If $x \in B$ then $i \leqslant x \wedge x \leqslant s$ and as $i \notin B$ we must have $i < x \leqslant s \Rightarrow x \in]i, s]$ so that $B \subseteq]i, s]$. This proves that $B =]i, s]$ a generalized interval.

C. $(\boldsymbol{i \in B \wedge s \notin B})$ If $x \in [i, s[$ then either $x = i \in B$ or $i < x < s \Rightarrow i < s \wedge i < x < s \underset{(4)}{\Rightarrow} x \in B$, proving that $[i, s[\subseteq B$. If $x \in B$ then $i \leqslant x \wedge x \leqslant s$ as $s \notin B$ we have $x < s \Rightarrow i \leqslant x < s \Rightarrow x \in [i, s[\Rightarrow B \subseteq [i, s[$. This proves that $B = [i, s[$ a generalized interval.

D. $(\boldsymbol{i, s \notin B})$ If $x \in]i, s[$ then $i < x < s \Rightarrow i < s \wedge i < x < s \underset{(4)}{\Rightarrow} x \in B$ proving that $]i, s[\subseteq B$. On the other hand if $x \in B$ then $i \leqslant x \leqslant s$ and as $i, s \notin B$ we must have $i < x < s \Rightarrow x \in]i, s[\Rightarrow B \subseteq]i, s[$. So $B =]i, s[$ a generalized interval.

iii. **($\boldsymbol{B}$ has only $\boldsymbol{a}$ upper bound)** then $s = \sup(B)$ exists by conditional completion. We have then the following cases

A. $(\boldsymbol{s \in B})$ If $x \in]-\infty, s]$ then $x \leqslant s$ and either $x = s \in B$ or $x < s$ and as there is no lower bound there exists a $b \in B$ such that $b \leqslant x < s$ so that either $x = b \in B$ or $b < x < s \underset{(4)}{\Rightarrow} x \in B$. We conclude thus that $]-\infty, s] \subseteq B$. If $x \in B$ then $x \leqslant s \Rightarrow x \in]-\infty, s] \Rightarrow B \subseteq]-\infty, s]$. So we conclude that $B =]-\infty, s]$ a generalized interval.

B. $(\boldsymbol{s \notin B})$ If $x \in]-\infty, s[$ then $x < s$ and as there is no lower bound there exists a $b \in B$ with $b \leqslant x$. Then either $x = b \in B$ or $b < x < s \underset{(4)}{\Rightarrow} x \in B$ proving that $]-\infty, s[\subseteq B$. If $x \in B \Rightarrow x \leqslant s$ and as $s \notin B$ we have $x < s \Rightarrow x \in]-\infty, s[\Rightarrow B \subseteq]-\infty, s[$. So we conclude that $B =]-\infty, s[$ a generalized interval.

iv. **(B has only a lower bound)** then $i = \inf(B)$ exists by conditional completion. We have then the following cases

A. **($i \in B$)** If $x \in [i, \infty[$ then $i \leqslant x$ and either $x = i \in B$ or $i < x$ and as there is no upper bound there exists a $b \in B$ with $i < x \leqslant b$ so that either $x = b \in B$ or $i < x < b \underset{(4)}{\Rightarrow} x \in B$. So we conclude that $[i, \infty[\subseteq B$. If $x \in B$ then $i \leqslant x \Rightarrow x \in [i, \infty[\Rightarrow B \subseteq [i, \infty[$. This finally means that $B = [i, \infty[$ a generalized interval.

B. **($i \notin B$)** If $x \in]i, \infty[$ then $i < x$ and as there is no upper bound there exists a $b \in B$ with $i < x \leqslant b$ so that either $x = b \in B$ or $i < x < b \underset{(4)}{\Rightarrow} x \in B$, thus we have $]i, \infty[\subseteq B$. If $x \in B$ then $i \leqslant x$ and as $i \notin B$ we must have $i < x \Rightarrow x \in]i, \infty[\Rightarrow B \subseteq]i, \infty[$. So we have $B =]i, \infty[$ a generalized interval □

2.3.2 Well-ordered classes

Definition 2.181. *A partially ordered class $\langle A, \leqslant \rangle$ is **well-ordered** if every nonempty subclass of A has a least element.*

Theorem 2.182. *If $\langle A, \leqslant \rangle$ is a partially ordered class and $B \subseteq A$ then we have the following for $\langle B, \leqslant \rangle$ (see 2.139)*

1. *If $\langle A, \leqslant \rangle$ is fully ordered then $\langle B, \leqslant \rangle$ is full ordered*
2. *If $\langle A, \leqslant \rangle$ is well-ordered then $\langle B, \leqslant \rangle$ is well-ordered*

Proof.

1. If $x, y \in B \Rightarrow x, y \in A \Rightarrow x \leqslant y \vee y \leqslant x \Rightarrow x \leqslant_{|B} y \vee y \leqslant_{|B} x$
2. If $C \subseteq B$ is a nonempty set then $C \subseteq A$ then there exists a least element c of C. So $c \in C$ and $\forall x \in C \vdash c \leqslant x \Rightarrow c \in C$ and $\forall x \in C \vdash c \leqslant_{|B} x$ □

Theorem 2.183. *A well-ordered class $\langle A, \leqslant \rangle$ is fully ordered*

Proof. If $x, y \in A \Rightarrow \{x, y\}$ is a nonempty subclass and thus it has a least element. We have then the following cases

1. x is the least element then $x \leqslant y$
2. y is the least element then $y \leqslant x$ □

Theorem 2.184. *A well-ordered class $\langle A, \leqslant \rangle$ is conditional complete.*

Proof. If B is a non empty subclass that is bounded above then $\upsilon(B) \neq \emptyset$ and thus $\upsilon(B)$ has a least element which is by definition $\sup(B)$ □

Definition 2.185. *Let $\langle A, \leqslant \rangle$ be a partially ordered class and $x, y \in B$ then y is a **immediate successor** of x if $x < y$ and there is no element $a \in A$ with $x < a < y$*

Theorem 2.186. *Let $\langle A, \leqslant \rangle$ be a well-ordered class then every element in A (with the exception of the greatest element) has a immediate successor*

Proof. If $x \in A$ and x is not the greatest element then $B = \{y \in A | y > x\}$ is nonempty and thus by well-ordering there exists a least element $b \in B$. Then $x < b$ and if $x < a < b$ we would have $a \in B$ and $a < b$ now b is the least element and thus $b \leqslant a \Rightarrow a < a \Rightarrow a \neq a$ a contradiction, □

Definition 2.187. *Let $\langle A, \leqslant \rangle$ be a partially ordered class and $B \subseteq A$ a subclass then B is a **section** of A iff*

$$\forall x \in A \vdash \text{if } y \in B \wedge x \leqslant y \Rightarrow x \in B$$

Theorem 2.188. *Let $\langle A, \leqslant \rangle$ be a well-ordered class then $B \subseteq A$ is a section of A iff $A = B$ or B is a initial segment of A*

Proof.

$\Rightarrow$

If B is a section of A then if $B = A$ we are done, so consider $B \neq A$ then $A \backslash B \neq \emptyset$. Because A is well-ordered $A \backslash B$ has a least element l we prove then that $B = S_l$

$$\begin{aligned} x \in S_l &\Rightarrow x \in A \wedge x < l \\ &\Rightarrow x \in B [\text{if } x \notin B \Rightarrow x \in A \backslash B \Rightarrow l \text{ is not the least element}] \\ x \in B &\Rightarrow \text{if } l \leqslant x \underset{B \text{ is a section}}{\Rightarrow} l \in B \text{ contradicting } l \in A \backslash B \Rightarrow x < l \Rightarrow x \in S_l \end{aligned}$$

$\Leftarrow$

If $A = B$ or B is a initial segment of A then either

1. $A = B$ then if $\forall x \in A \vdash y \in B \wedge x \leqslant y$ we have $x \in A$
2. $\exists b \in A \vdash B = S_b = \{x \in A | x < b\}$ so if $x \in A$ and there exists a $y \in B$ with $x \leqslant y \underset{y < b}{\Rightarrow} x < b \Rightarrow x \in B$ □

Theorem 2.189. (Transfinite Induction) *Let $\langle A, \leqslant \rangle$ be a well-ordered class and let $P(x)$ be a statement which is either false or true for a element $x \in A$ which satisfies*

$$\text{If } P(y) \text{ is true for every } y < x \text{ then } P(x) \text{ is true}$$

then $P(x)$ is true for every $x \in A$

Proof. Suppose $P(x)$ is not true for every $x \in A$ then $B = \{y \in A | P(x) \text{ is false}\}$ is not empty. As A is well-ordered there exists a least element b of B. Now if $x < b$ then if $x \in B \Rightarrow x < b \leqslant x \Rightarrow x < x$ a contradiction. So if $x < b$ then $x \notin B$ and thus $P(x)$ is true. By the hypothesis we have then that $P(b)$ is true which means $b \notin B$ contradicting the fact that b is the least element of B. So we must have that $P(x)$ is true for every $x \in A$. □

Theorem 2.190. *If $\langle A, \leqslant \rangle$ is a well-ordered class, $B \subseteq A$ and $f: A \rightarrow B$ a isomorphism then $\forall x \in A \vdash x \leqslant f(x)$*

Proof. We prove this by contradiction so assume that there exists a $x \in A$ such that $f(x) < x$. Then we have that $C = \{x \in A | f(x) < x\} \neq \emptyset$ and by well-ordering it has a least element c. As $c \in C$ we have $f(c) < c$ and so using the fact that f is a isomorphism we have by 2.155 that $f(f(c)) < f(c)$ so $f(c) \in C$ but as c is the least element we have $f(c) < c \leqslant f(c) \Rightarrow f(c) < f(c)$ a contradiction. So we must conclude that $C = \emptyset$ proving our theorem. □

Theorem 2.191. *Let $\langle A, \leqslant \rangle$ be a well-ordered class then there does not exists a isomorphism from A to a subclass of an initial segment of A.*

Proof.

We prove this again by contradiction. So assume that there exists a $a \in A$, a $B \subseteq S_a$ and a isomorphism $f: A \rightarrow B$. Then by the previous theorem we have $a \leqslant f(a) \Rightarrow \neg(f(a) < a)$ and thus $f(a) \notin S_a \Rightarrow f(a) \notin B$ but this contradicts the fact that $\text{range}(f) = B$. Hence the theorem must be true. □

As a initial segment is a subclass of itself so we have the following corollary

Corollary 2.192. *Let $\langle A, \leqslant \rangle$ be a well-ordered class then there does not exists a isomorphism from A to a initial segment of A.*

Theorem 2.193. *If $\langle A, \leqslant_A \rangle$, $\langle B, \leqslant_B \rangle$ be well-ordered classes then if A is isomorphic with an initial segment of B we have that B is not isomorphic with any subclass of A*

Proof. Let $b \in B$ and let $f: A \rightarrow S_b$ be a isomorphism from A to a initial segment of B. We proceed now by contradiction, so assume that there exists a $C \subseteq A$ and a isomorphism $g: B \rightarrow C$ then $g: B \rightarrow A$ is a injective increasing function hence $f \circ g: B \rightarrow S_b$ is a injective increasing function (see 2.46 and 2.151) and thus $f \circ g: B \rightarrow (f \circ g)(B) \subseteq S_b$ is a injective increasing function that by 2.159, 2.183 is a isomorphism. So we have a isomorphism of B to a subclass of one of its initial segments which by 2.191 is impossible. Proving the theorem by contradiction. □

Lemma 2.194. *Let $\langle A, \leqslant \rangle$ be a well-ordered and $a, b \in A$ with $a < b$ then S_a is a initial segment of S_b (using the induced order relation on S_b).*

Proof. First

$$\begin{aligned} x \in S_a &\Rightarrow x < a \\ &\underset{a<b}{\Rightarrow} x < b \\ &\Rightarrow x \in S_b \end{aligned}$$

so we have $S_a \subseteq S_b$. Now if $x \in S_b$ and there is a $y \in S_a$ such that $x \leqslant y \Rightarrow x \leqslant y < a \Rightarrow x < a \Rightarrow x \in S_a$. So S_a is a section of S_b and as $a \notin S_a \wedge a \in S_b$ we have $S_a \neq S_b$ so that using 2.188 S_a is a initial segment of S_b. □

Theorem 2.195. *Let $\langle A, \leqslant_A \rangle$, $\langle B, \leqslant_B \rangle$ be well-ordered classes then exactly one of the following cases holds*

1. *A is isomorphic with B*

2. *A is isomorphic with an initial segment of B*
3. *B is isomorphic with an initial segment of A*

Proof. Let's define $C = \{x \in A | \exists y \in B \vdash S_x \cong S_y\}$ we have then the following

$$x \in C \;\Rightarrow\; \exists! y \vdash S_x \approx S_y \,(\text{there is only on } y \text{ such that } S_x \cong S_y)$$

Suppose that given $x \in C$ there exists y, y' with $y \neq y'$ such that $S_x \cong S_y$ and $S_x \cong S_{y'}$ giving $S_y \cong S_{y'}$ and $S_{y'} \cong S_y$(see 2.158) then we have as B is well-ordered and thus fully ordered the following cases

1. $y \leqslant_B y' \Rightarrow y <_B y'$ so by the previous lemma we have that S_y is a initial segment of $S_{y'}$ and then from $S_{y'} \cong S_y$ we have that $S_{y'}$ is isomorphic to a initial segment of itself which is forbidden by 2.192.
2. $y' \leqslant_B y \Rightarrow y' <_B y$ so by the previous lemma we have that $S_{y'}$ is a initial segment of S_y and then from $S_y \cong S_{y'}$ we have that S_y is isomorphic to a initial segment of itself which is forbidden by 2.192.

So we must conclude that $y = y'$ proving our assessment.

Using the above we can define a function $F: C \rightarrow B$ where $F = \{(x, y) | S_x \approx S_y\}$ so if $x \in C$ then $S_x \cong S_{F(x)}$. We prove now that $F: C \rightarrow D$ where $D = \text{range}(F) \subseteq B$ is a isomorphism.

1. **(injective)** If $F(x) = F(x') = d$ then $S_x \cong S_d$ and $S_{x'} \cong S_d \Rightarrow S_x \cong S_{x'}$ now by fully ordering we have if $x \neq x'$

   a. $x \leqslant_A x' \Rightarrow x <_A x'$ so S_x is a initial segment of $S_{x'}$ which by 2.192 conflicts with $S_x \cong S_{x'}$

   b. $x' \leqslant_A x \Rightarrow x' <_A x$ so $S_{x'}$ is a initial segment of S_x which by 2.192 conflicts with $S_{x'} \cong S_x$

   proving that $x \neq x'$ leads to a contradiction, so we must have $x = x'$.

2. **(surjective)** This is trivial as $D = \text{range}(F)$

3. **(increasing)** Suppose $x \leqslant_A y$ then $S_x \cong S_{F(x)}$ and $S_y \cong S_{F(y)}$. Proceed now by contradiction so assume $\neg(F(x) \leqslant_B F(y))$ or $F(y) <_B F(x)$ then $S_{F(y)}$ is a initial segment of $S_{F(x)}$ and from $x \leqslant_A y$ we have $S_x \subseteq S_y$ which proves that

   a. S_y is isomorphic with $S_{F(y)}$ a initial segment of $S_{F(x)}$

   b. $S_{F(x)}$ is isomorphic with S_x a subclass of S_y

   which by 2.193 is forbidden. So we are left with $F(x) \leqslant_B F(y)$ and thus F is increasing.

From 1,2 and 3 and 2.159 we have that $F: C \rightarrow D$ is a isomorphism.

Next we show that ***C* is a section of *A***. So if $x \in A$ and there is a $c \in C$ such that $x \leqslant_A c$ then $S_c \cong S_{F(c)}$ and thus there exists a isomorphism $g: S_c \rightarrow S_{F(c)}$ we prove now that

$$g_{|S_x}: S_x \rightarrow S_{g(x)} \quad \text{ is } a \text{ isomorphism}$$

First $x \leqslant_A c \Rightarrow S_x \subseteq S_c$ and if $y \in S_x \Rightarrow y <_A x \underset{g \text{ is } a \text{ isomorphism}}{\Rightarrow} g(y) <_B g(x) \Rightarrow g(y) \in S_{g(x)}$ and thus $g_{|S_x}: S_x \rightarrow S_{g(x)}$ is a function. We prove now that it is a isomorphism

1. **(injective)** If $g_{|S_x}(y) = g_{|S_x}(y') \Rightarrow g(y) = g(y') \underset{g \text{ is } a \text{ isomorphism}}{\Rightarrow} y = y'$
2. **(surjective)** If $y \in S_{g(x)} \Rightarrow y <_B g(x) \in S_{F(c)} \Rightarrow g(x) <_B F(c) \Rightarrow y <_B F(c) \underset{g \text{ is surjective}}{\Rightarrow} \exists r \in S_c$ such that $g(r) = y$. Now if $x \leqslant_A r$ then $g(x) \leqslant_B g(r) = y$ contradicting $y <_B g(x)$ so we must have $r <_A x \Rightarrow r \in S_x$ and $y = g(r) = g_{|S_x}(r) \Rightarrow \text{range}(g_{|S_x}) = S_{g(x)}$
3. **(increasing)** If $s, r \in S_x$ and $s \leqslant_A r \Rightarrow g_{|S_x}(s) = g(s) \leqslant_B g(r) = g_{|S_x}(r)$

So by using 2.159 we have that $g_{|S_x}$ is a isomorphism. Thus $S_x \cong S_{g(x)}$ and thus $x \in C$ proving that **C is a section of A.**

Next we prove that **D is a section of B**, so if $y \in B$ is such that there is a $d \in D$ with $y \leqslant_B d$. Then as $d \in D = \text{range}(F)$ there exists a $c \in C \vdash F(c) = d$ so that $S_c \cong S_d \Rightarrow S_d \approx S_c$. So there exists a isomorphism $g: S_d \rightarrow S_c$. Now from $y \leqslant_B d$ we have that $S_y \subseteq S_d$ and if $x \in S_y \Rightarrow x <_B y \underset{g \text{ is } a \text{ isomorphism}}{\Rightarrow} g(x) <_A g(y) \Rightarrow g(x) \in S_{g(y)}$ so $g_{|S_y}: S_y \rightarrow S_{g(y)}$ is a function, we prove now that

$$g_{|S_y}: S_y \rightarrow S_{g(y)} \quad \text{is } a \text{ isomorphism}$$

1. **(injective)** If $g_{|S_y}(x) = g_{|S_y}(x') \Rightarrow g(x) = g(x') \underset{g \text{ is injective}}{\Rightarrow} x = x'$
2. **(surjective)** If $z \in S_{g(y)} \Rightarrow z <_A g(y) \in S_c \Rightarrow g(y) <_A c \Rightarrow z <_A c \Rightarrow z \in S_c$ so there exists a $x \in S_d$ such that $g(x) = z$. Now if $y \leqslant_B x \Rightarrow g(y) \leqslant_A g(x) = z \Rightarrow g(y) \leqslant_A z \Rightarrow z \notin S_{g(y)}$ contradicting $z \in S_{g(y)}$ so we must have $x <_B y \Rightarrow x \in S_y$ and $g_{|S_y}(x) = g(x) = z \Rightarrow \text{range}(g_{|S_y}) = S_{g(y)}$
3. **(increasing)** If $s, r \in S_y$ and $s \leqslant_B r \Rightarrow g_{|S_y}(s) = g(s) \leqslant_A g(r) = g_{|S_y}(r)$

So using 2.159 we have that $g_{|S_y}: S_y \rightarrow S_{g(y)}$ is a isomorphism. Thus $S_y \cong S_{g(y)}$ and thus as $g(y) \in S_c \subseteq A$ we have $g(y) \in C$ and $F(g(y)) = y \Rightarrow y \in D$ proving that **D is a section of B.**

We have thus proved that $F: C \rightarrow D$ is a isomorphism where C is a section of A and D is a section of B. Using 2.188 we have then the following cases for the isomorphism $F: C \rightarrow D$

1. $C = A \wedge D = B$ then A is isomorphic with B
2. $C = A \wedge D$ is a initial segment of B then A is isomorphic with a initial segment of B
3. C is a initial segment of $A \wedge D = B$ then B is isomorphic with a initial segment of B
4. C is a initial segment of A and D is a initial segment of B. In this case there exists a $a \in A$ and a $b \in B$ so that $C = S_a$ and $D = S_b$ and $S_a \cong S_b$ but then $a \in C \Rightarrow a \in S_a \Rightarrow a < a$ a contradiction. So this case does not occur.

We conclude that or theorem is proved if we prove that 1,2 and 3 can not occur at the same time but this follows from 2.193. □

Corollary 2.196. *Let $\langle A, \leqslant_A \rangle$ be a well-ordered class then every subclass of A is isomorphic with A or an initial segment of A*

Proof. If $B \subseteq A$ then $\langle B, \leqslant \rangle$ is a well-ordered class and using the previous theorem we have the following exclusive cases

1. A is isomorphic with B
2. B is isomorphic with a initial segment of A then as $B \cong B$ we use 2.193 to reach a contradiction. So this can not occur.
3. B is isomorphic with a initial segment of A □

2.3.3 The Axiom of Choice

Definition 2.197. *Let A be a set then a **choice** function for A is a function c: $\mathcal{P}'(A) \to A$ such that $\forall B \in \mathcal{P}'(A) \vdash c(B) \in B$ (see 1.72*

We can then state the axiom of choice as follows

Axiom 2.198. (Axiom of Choice) *Every set has a choice function.*

Using the Axiom of Choice we can prove the opposite from 2.45

Theorem 2.199. *If $f: A \to B$ is a surjective function then there exists a function $g: B \to A$ such that $f \circ g = 1_B$*

Proof. If $f: A \to B$ is surjective. Then $\forall y \in B$ we have that $f^{-1}(\{y\})$ is a non empty subset of $A \Rightarrow f^{-1}(\{y\}) \in \mathcal{P}'(A)$. By the axiom of choice there exists a choice function $c: \mathcal{P}'(A) \to A$. Define now $g: B \to A$ such that $g(y) = c(f^{-1}(y))$. Then if $y \in B$ we have $g(y) = c(f^{-1}(y)) \in f^{-1}(\{y\}) \Rightarrow f(g(y)) = y$ and thus $f \circ g = 1_B$. we prove now that g is injective. If $g(y) = g(y')$ then we have $f(g(y)) = f(g(y')) \Rightarrow 1_B(y) = 1_B(y') \Rightarrow y = y'$. □

Using the above theorem and 2.45 we have

Theorem 2.200. *A function $f: A \to B$ is surjective if and only if there exists a function $g: B \to A$ such that $f \circ g = 1_B$*

We have the following equivalent definitions of the axiom of choice

Theorem 2.201. *The following are equivalent*

1. *The Axiom of Choice*
2. *Let $\mathcal{A}$ be a set of sets such that $A \in \mathcal{A} \Rightarrow A \neq \emptyset$ and if $A, B \in \mathcal{A}$ with $A \neq B$ then $A \bigcap B = \emptyset$ (in other words $\mathcal{A}$ is a set of nonempty mutually disjoint sets) then there exists a set C (called the choice set for $\mathcal{A}$) such that $\forall A \in \mathcal{A} \vdash \exists! a \in A$ so that $a \in C$ and if $a \in C$ then $\exists A \in \mathcal{A}$ such that $a \in A$ (in other words a choice set consists of exactly one element of every element of $\mathcal{A}$)*
3. *If $\{A_i\}_{i \in I}$ is a family of nonempty set then $\exists c: I \to \bigcup_{i \in I} A_i$ such that $\forall i \in I \vdash c(i) \in A_i$ (in other words $\prod_{i \in I} A_i \neq \emptyset$)*

Proof.

$1 \Rightarrow 2$

So let $\mathcal{A}$ be a set of nonempty mutually disjoint sets and let $\mathbb{A} = \bigcup_{X \in \mathcal{A}} X$ then we have $\mathcal{A} \subseteq \mathcal{P}'(\mathbb{A})$, now by the axiom of choice there is a function $c: \mathcal{P}'(\mathcal{A}) \to \mathbb{A}$. Take then $C = c(\mathcal{A})$ then if $A \in \mathcal{A} \subseteq \mathcal{P}'(\mathbb{A}) \Rightarrow c(A) \in A$ and also $c(A) \in C$ so C consist of a element of all sets in $\mathcal{A}$. If now $A \in \mathcal{A}$ and $X \in \mathcal{A} \vdash c(X) \in A \underset{c(X) \in X}{\Rightarrow} X \bigcap A \neq \emptyset \Rightarrow X = A$ (as we have mutually disjoint sets). So if there exists a $A \in \mathcal{A}$ with a, $a' \in A$ and $a, a' \in C = c(\mathcal{A})$ then $\exists X, X' \in \mathcal{A} \vdash c(X) = a \in A, c(X') = a' \in A \Rightarrow X = A = X' \underset{X = X' = A \wedge (X, a) \in c, (X', a') \in c}{\Rightarrow} (A, a), (A, a') \in c \underset{c \text{ is a function}}{\Rightarrow} a = a'$ proving that C is indeed the choice set for $\mathcal{A}$

$2 \Rightarrow 1$

Let A be a set and let $B \subseteq A$ be a nonempty subset of A. Form then $P_B = \{(B, x) | x \in B\}$, Then we have that P_B is nonempty as B is nonempty. Also if $z \in P_B \bigcap P_{B'} \Rightarrow z = (B, x) \wedge x \in B \wedge z = (B', x') \wedge x' \in B' \Rightarrow B = B' \Rightarrow P_B = P_{B'}$. So if $P_B \neq P_{B'}$ then $P_B \bigcap P_{B'} = \emptyset$. From this it follows that $\mathcal{A} = \{P_B | B \subseteq A \wedge B \neq \emptyset\}$ is set of nonempty mutually disjoint sets (note that $P_B \subseteq \mathcal{P}(A) \times A$ and thus $\mathcal{A} \subseteq \mathcal{P}(\mathcal{P}(A) \times A)$ is indeed a set). From the hypothesis we have then the existence of a choice set C. If $B \in \mathcal{P}'(A)$ then $P_B \in \mathcal{A} \Rightarrow \exists (B, b) \in P_B \vdash (B, b) \in C$. If $c \in C$ then $\exists B \in \mathcal{P}'(A) \vdash c \in P_B \Rightarrow \exists b \in B \subseteq A \vdash (B, b) = c \Rightarrow C \subseteq \mathcal{P}'(A) \times A$. Also if $(B, b) \in C \Rightarrow \exists P_{B'} \vdash (B, b) \in P_{B'} \Rightarrow B = B' \wedge b \in B' \Rightarrow b \in B$. So if $(B, b), (B, b') \in C \Rightarrow b$, $b' \in B \Rightarrow (B, b), (B, b') \in P_B \underset{C \text{ is a choice set}}{\Rightarrow} (B, b) = (B, b') \Rightarrow b = b'$. So we have proved that $C: \mathcal{P}'(A) \to A$ is a function such that $\forall B \in \mathcal{P}'(A) \vdash C(B) \in B$. So C is the choice function we would search for.

$1 \Rightarrow 3$

If $\{A_i\}_{i \in I}$ is a family of nonempty sets, form then $A = \bigcup_{i \in I} A_i$ by the axiom of choice there exists a $c: \mathcal{P}'(A) \to A$ such that $\forall B \in \mathcal{P}'(A)$ we have $c(B) \in B$. Define then the function $x: I \to \mathcal{A}$ with $x = \{(i, y) | i \in I \wedge y = c(A_i)|\} \subseteq I \times A$ [then $\text{dom}(x) = I$ and $(i, y), (i, y') \in x \Rightarrow y = c(A_i) = y' \Rightarrow y = y'$ so x is a function and $\forall i \in I \vdash x(i) = c(A_i) \in A_i$] and thus $x \in \prod_{i \in I} A_i$

$3 \Rightarrow 1$

If A is a set define then the graph $\mathcal{B} = \{(B, b) | B \in \mathcal{P}'(A) \wedge b \in B\}$. We have then $\text{dom}(\mathcal{B}) \subseteq \mathcal{P}'(A)$ and if $B \in \mathcal{P}'(A)$ there exists a $b \in B$ and thus $(B, b) \in \mathcal{B}$ proving that $\text{dom}(\mathcal{B}) = \mathcal{P}'(A)$. This defines then the family $\{\mathcal{B}_B\}_{B \in \mathcal{P}'(A)}$. Now given $B \in \mathcal{P}'(A)$ we have

$$\begin{aligned} x \in B &\Leftrightarrow B \in \mathcal{P}'(A) \wedge x \in B \\ &\Leftrightarrow (B, x) \in \mathcal{B} \\ &\Leftrightarrow x \in \{y | (B, y) \in \mathcal{B}\} \\ &\Leftrightarrow x \in \mathcal{B}_B \end{aligned}$$

proving that $\mathcal{B}_B = B$. So $\bigcup_{B \in \mathcal{P}'(A)} \mathcal{B}_B \subseteq A$ and as $x \in A \Rightarrow \{x\} \in \mathcal{P}'(A) \Rightarrow x \in \{x\} = \mathcal{B}_{\{x\}} \Rightarrow x \in \bigcup_{B \in \mathcal{P}'(A)} \mathcal{B}_B$ so $A = \bigcup_{B \in \mathcal{P}'(A)} \mathcal{B}_B$ and by the hypothesis there exists a $c: \mathcal{P}'(A) \to A$ with $\forall B \in \mathcal{P}'(A) \vdash c(B) \in B$ and c is the sought for choice function. □

Corollary 2.202. *Let A and B be sets and assume that $\forall x \in A$ there exists a $y \in B$ such that $P(x, y)$ is true then there exists a function $f: A \to B$ such that $\forall x \in A$ we have $P(f(x))$ is true*

Proof. Define the graph $G = \{(x, y) \in A \times B | P(x, y)\}$ then this defines (see 1.84) a family $\{F_a\}_{a \in A}$ with $F_a = \{y \in | P(a, y)\} \neq \emptyset$ so using the equivalences of the Axiom of Choice (see 2.201) there exists a function $f: A \rightarrow \bigcup_{a \in A} F_a$ with $\forall a \in A$ that $f(a) \in F_a \Rightarrow P(a, f(a))$ is true. □

As a direct application of the Axiom of Choice we prove the following

Theorem 2.203. *Let $\langle A, \leqslant \rangle$ be a partially ordered set such that*

1. *A has a least element p*
2. *Every chain of A has a supremum*

then there is a element $x \in A$ which has no immediate successor

Proof. We will prove this by contradiction. So assume that $\forall x \in A$ there exists a immediate successor, we will then derive a contradiction from this.

First $\forall x \in A$ define $T_x = \{y: y \text{ is a immediate successor of } x\}$ then $T_x \neq \emptyset \in \mathcal{P}'(A)$. By the axiom of choice there exists a choice function from $c: \mathcal{P}'(A) \rightarrow A$ such that $\forall A \in \mathcal{P}'(A) \vDash c(A) \in A$. This lets us define a function $f: A \rightarrow A$ where $f = \{(x, y): x \in A \wedge y = c(T_x)\} \subseteq A \times A$ [if $x \in A \Rightarrow T_x \neq \emptyset \Rightarrow (x, c(T_x)) \in f \Rightarrow \operatorname{dom}(f) = A$ and if $(x, y), (x, y') \in f \Rightarrow y = c(T_x) = y' \Rightarrow y = y'$ thus $f: A \rightarrow A$ is indeed a function] and if $x \in A \Rightarrow f(x) \in T_x$ so $f(x)$ is a immediate successor of x. So we have defined a function

$$f: A \rightarrow A \quad \forall x \in A \vDash f(x) \text{ is a immediate successor of } x$$

Given the least element p of A we define the concept of a **p-sequence** as follows

Definition 2.204. *A subset $B \subseteq A$ is called a **p-sequence** iff*

1. *$p \in B$*
2. *If $x \in B \Rightarrow f(x) \in B$*
3. *If C is a chain of B (and thus of A) then $\sup(C) \in B$*

Note that A is trivially a p-sequence. We have now the following lemma

Lemma 2.205. *Every intersection of p-sequences is a p-sequence*

Proof. If $\mathcal{A}$ is a collection of $p-$sequences then

1. $\forall A \in \mathcal{A} \vDash p \in A \Rightarrow p \in \bigcap_{A \in \mathcal{A}} A$
2. If $x \in \bigcap_{A \in \mathcal{A}} A \Rightarrow \forall A \in \mathcal{A} \vDash x \in A \Rightarrow \forall A \in \mathcal{A} \vDash f(x) \in A \Rightarrow f(x) \in \bigcap_{A \in \mathcal{A}} A$
3. If C is a chain in $\bigcap_{A \in \mathcal{A}} A$ then every element in C is comparable and $C \subseteq \bigcap_{A \in A} A \Rightarrow \forall A \in \mathcal{A} \vDash C \subseteq A \Rightarrow \forall A \in \mathcal{A} \vDash C \text{ is a chain in } A \Rightarrow \forall A \in \mathcal{A} \vDash \sup(C) \in A \Rightarrow \sup(C) \in \bigcap_{A \in \mathcal{A}} A$

So we have that $\bigcap_{A \in \mathcal{A}} A$ is a p-sequence □

Take now $\mathcal{A} = \{B \in \mathcal{P}(A): B \,\text{is}\, ap\text{−sequence}\}$ which is nonempty as it contains A and define the intersection of all p-sequences

Theorem 2.206. *Let $\{A_i\}_{i \in I}$ be a family of sets and $b: J \to I$ a bijection then*

$$\bigcap_{i \in I} A_i = \bigcap_{j \in J} A_{b(j)}$$
$$\bigcup_{i \in I} A_i = \bigcup_{j \in J} A_{b(j)}$$

Proof. First we prove that

□

$$P = \bigcap_{B \in \mathcal{A}} B$$

then P is a p-sequence and (as $p \in P$ we have $P \neq \emptyset$) and if B is a p-sequence then $P \subseteq B$. Next we define the concept of **select** elements in P.

Definition 2.207. *A element $x \in P$ is called **select** if it is comparable with every element y of P*

We have then the following lemma

Lemma 2.208. *Suppose x is **select** then if $y \in P$ with $y < x$ we have $f(y) \leqslant x$*

Proof. If $y \in P$ then as P is a p-sequence we have by (2) that $f(y) \in P$. Now x is select so we have either $f(y) \leqslant x$ or $x < f(y)$. If $x < f(y)$ then from $y < x$ we have $y < x < f(y)$ contradicting that $f(y)$ is the immediate successor of y, so we must have $f(y) \leqslant x$ □

Next we prove the following lemma

Lemma 2.209. *Suppose x is select then $B_x = \{y \in P | y \leqslant x \,\text{or}\, f(x) \leqslant y\}$ is a p-sequence.*

Proof.

1. Since p is the least element of A we have $p \leqslant x \Rightarrow p \in B_x$
2. Suppose $y \in B_x$ then $y \leqslant x$ or $y \geqslant f(x)$ lets break this in the following three cases
   a. $y < x$. Then by the previous lemma we have $f(y) \leqslant x \Rightarrow f(y) \in B_x$
   b. $y = x$. Then $f(y) = f(x)$ thus $f(x) \leqslant f(y) \Rightarrow f(y) \in B_x$
   c. $y \geqslant f(x)$ Then as $f(y)$ is the immediate successor of y we have $y < f(y) \Rightarrow f(x) < f(y) \Rightarrow f(x) \leqslant f(y) \Rightarrow f(y) \in B_x$

   So in all cases we have $f(y) \in B_x$
3. If C is a chain in B_x (and thus in A) let $s = \sup(C)$ then we have the following excluding cases
   a. $\exists y \in C \vdash f(x) \leqslant y \underset{y \leqslant s}{\Rightarrow} f(x) \leqslant s \Rightarrow s \in B_x$

b. $\forall y \in C \vDash \neg(f(x) \leqslant y)$. Now $\forall y \in C$ as $y \in C \subseteq B_x$ we have $y \leqslant x$ or $f(x) \leqslant y$, as the last contradicts $\neg(f(x) \leqslant y)$, we have always $y \leqslant x$ and thus x is a upper bound of C. As s is the least upper bound we must have $s \leqslant x \Rightarrow s \in B_x$

So in all valid cases $s \in B_x$ □

We can easily derive the following corollary.

Corollary 2.210. *If x is select then $\forall y \in P$ we have $y \leqslant x$ or $f(x) \leqslant y$*

Proof. As $P \subseteq B_x$ (P is the intersection of all the p-sequences) and by definition of B_x we have $B_x \subseteq P \Rightarrow P = B_x$ □

We can use this to prove the following lemma

Lemma 2.211. *The set of all select elements $\{x \in P | x \,\mathrm{is\,select}\}$ is a p-sequence*

Proof.

1. p is select because $\forall x \in P \subseteq A$ we have $p \leqslant x$ so it is comparable with every element of p
2. If x is select then, by the previous corollary, $\forall y \in P$ we have either

   a. $y \leqslant x$. Then as $x < f(x) \Rightarrow y < f(x) \Rightarrow y \leqslant f(x)$

   b. $f(x \leqslant y)$

   So $f(x)$ is comparable with every element of P and $f(x)$ is thus select.
3. Let C be a chain of select elements. Then as C is a chain in A we have that $s = \sup(C)$ exists. Then $\forall y \in P$ we have the following possibilities for C

   a. $\exists x \in C \vdash y \leqslant x$ then as $x \leqslant s$ we have $y \leqslant s$

   b. $\forall x \in C \vdash \neg(y \leqslant x)$ then as $y \in P$ we have $\forall x \in C$ either $y \leqslant x$ (contradicted by $\neg(y \leqslant x)$) or $f(x) \leqslant y$ so $\forall x \in C$ we have $f(x) \leqslant y \underset{x < f(x)}{\Rightarrow} x \leqslant y$ and thus y is a upper bound of C. Ans as s is the least upper bound of S we have $s \leqslant y$

   So s is comparable with every $y \in P$ and thus s is select. □

Corollary 2.212. *P is fully-ordered*

Proof. $S = \{x \in P | x \,\mathrm{is\,select}\} \subseteq P$ and by the previous lemma we have that S is a p-sequence thus $P \subseteq S$ so $P = S = \{x \in P | x \,\mathrm{is\,select}\}$. Or every element in P is select and thus comparable with every other element so P is fully-ordered. □

Now for the finally proof, we show that this corollary causes a contradiction. As P is by the corollary a chain we have that $s = \sup(P)$ exists by the hypothesis. But as P is a p-sequence we find by (2) that $f(s) \in P$ hence $f(s) \leqslant s$ but as $f(s)$ is the successor or s we have $s < f(s) \Rightarrow s < s$ a contradiction. So we have derived a contradiction and the theorem must be true. □

We show now that the following theorem follows from the Axiom of Choice (via 2.203)

Theorem 2.213. (Hausdorff's Maximal Principle) *Every partially ordered set $\langle A, \leqslant\rangle$ has a maximal chain C (see 2.161 if D is a chain and if $C \subseteq D \Rightarrow C = D$)*

Proof. Define $\mathcal{C} = \{B \in \mathcal{P}(A) | B \,\mathrm{is}\, a \,\mathrm{chain}\}$ be the set [is $\subseteq \mathcal{P}(A)$ and applying 1.69 and 1.64 gives $\mathcal{C}$ is a set] of all chains in A and order $\mathcal{C}$ by inclusion (see 2.133) so $\langle \mathcal{C}, \subseteq\rangle$ is a partially ordered set. Notice that $\mathcal{C}$ **has a least element** namely $\emptyset$. Let now $\mathcal{D} \subseteq \mathcal{C}$ be a chain in $\mathcal{C}$ and define

$$K = \bigcup_{C \in \mathcal{D}} C$$

then we prove that K is a chain in A and thus $K \in \mathcal{C}$. Indeed if $x, y \in K$ then $\exists C_1$, $C_2 \in \mathcal{D}$ such that $x \in C_1$ and $y \in C_2$. Now as $\mathcal{D}$ is a chain in $\mathcal{C}$ we have either

1. $C_1 \subseteq C_2 \Rightarrow x, y \in C_2 \underset{C_2 \,\mathrm{is}\, a \,\mathrm{chain\,in}\, A}{\Rightarrow} x$ and y are comparable
2. $C_2 \subseteq C_1 \Rightarrow x, y \in C_1 \underset{C_1 \,\mathrm{is}\, a \,\mathrm{chain\,in}\, A}{\Rightarrow} x$ and y are comparable

Theorem 2.214. *Let $\{A_i\}_{i \in I}$ be a family of sets and $b: J \rightarrow I$ a bijection then*

$$\bigcap_{i \in I} A_i = \bigcap_{j \in J} A_{b(j)}$$
$$\bigcup_{i \in I} A_i = \bigcup_{j \in J} A_{b(j)}$$

Proof. First we prove that

□

proving that K is a chain. Using 2.168 we have that $\boldsymbol{K = \sup(\mathcal{D})}$ exists. So we can apply 2.203 (which is derived from the Axiom of Choice) to find a element $C \in \mathcal{C}$ which has no immediate successor. C is (by definition of C) a chain. We prove now by contradiction that C is maximal. If there exists a chain D in A with $C \subseteq D$ and $D \neq C$ then $\exists d \in D \vdash d \notin C$. Then $C \bigcup \{d\}$ is a chain for if x, $y \in C \bigcup \{d\} \subseteq D$ we have as D is a chain that x, y are comparable. Now if there is a set X with $C \subset X \subset C \bigcup \{d\}$ then $X = X \bigcup (X \setminus C) \subseteq C \bigcup \{d\} \Rightarrow (X \setminus C) \subseteq C \bigcup \{d\}$ and if $x \in (X \setminus C) \Rightarrow x \notin C \wedge x \in C \bigcup \{d\} \Rightarrow x = d$ so $(X \setminus C) = \{d\}$ so $X = C \bigcup \{d\}$ contradicting $X \subset C \bigcup \{d\}$. So for every chain D in A with $C \subseteq D$ we must have $C = D$ proving maximality of C. □

Definition 2.215. *A partially ordered set $\langle A, \leqslant\rangle$ is said to be* ***inductive*** *when every chain in A has an upper bound.*

Lemma 2.216. (Zorn's Lemma) *Every inductive set has at least one maximal element. In other words, if $\langle A, \leqslant\rangle$ is a partially ordered set such that every chain in A has a upper bound, then A has a maximal element.*

We prove now that from the Hausdorff's maximal theorem follows directly Zorn's lemma, without using either 2.201 or 2.203 ,so it does not depends indirectly on the Axiom of Choice, but directly on Hausdorff's maximal theorem.

Theorem 2.217. *Hausdorff's maximal theorem implies* Zorn′*s lemma*

Proof. Let $\langle A, \leqslant\rangle$ be a partially ordered set such that every chain in A has a upper bound. By the Hausdorff's Maximal Principle (see 2.213) there exists a maximal chain C in A, by the hypothesis there exists a upper bound u for C. We prove now by contradiction that u is a maximal element of A. So assume that u is not a maximal element of A then there exists a $x \in A \vdash u < x$. If now $x \in C \underset{u\,\mathrm{is}\,a\,\mathrm{upper\,bound}}{\Rightarrow} x \leqslant u < x \Rightarrow x < x$ a contradiction so we must have that $x \notin C$. Now if $r, s \in C \bigcup \{x\}$ then we have the following cases

1. $r = s = x \underset{\text{reflexitivity}}{\Rightarrow} r \leqslant s$
2. $r = x, s \in C \Rightarrow s \leqslant u < x = r \Rightarrow s \leqslant r$
3. $r \in C, s = x \Rightarrow r \leqslant u < x = s \Rightarrow r \leqslant s$
4. $r, s \in C \underset{C\,\mathrm{is}\,a\,\mathrm{chain}}{\Rightarrow} s \leqslant r$ or $r \leqslant s$

proving that $C \bigcup \{x\}$ is a chain with (because $x \notin C$) $C \subset C \bigcup \{x\}$ contradicting that C is a maximal chain. So we must have that u is a maximal element of A. □

We state now the Well-Ordering Theorem

Theorem 2.218. (Well-Ordering Theorem) *Given a set A then there exists a order relation $\leqslant$ on A such that $\langle A, \leqslant\rangle$ forms a well-ordered set.*

We prove then that from the Zorn's lemma follows directly the Well-Ordering Theorem without using either 2.201,2.213 or 2.203 so it depends directly on Zorn's Lemma.

Theorem 2.219. *Zorn's Lemma implies the Well-Ordering Theorem*

Proof. Let A be a arbitrary set and let $B \subseteq A$ a subset. A relation G on B is a subclass of the set $B \times B$ (see 1.79) and we have by 1.64 that G is a set. So the pair (B, G) is a element (see 1.39) we define then the class $\mathcal{A} = \{(B, G) | B \in \mathcal{P}(A) \wedge G\,\mathrm{is}\,a\,\mathrm{order\,relation\,on}\,B\,\mathrm{so\,that}\langle B, G\rangle\,\mathrm{is\,well{-}ordered}\}$. We define now $\prec \subseteq \mathcal{A} \times \mathcal{A}$ by

$$(B, G) \prec (B', G') \Leftrightarrow B \subseteq B'\,\mathrm{and}\,G \subseteq G'\,\mathrm{and}\,[x \in B \wedge y \in B' \backslash B \Rightarrow (x, y) \in G'] \tag{2.20}$$

Lemma 2.220. $\langle \mathcal{A}, \prec\rangle$ *forms a order relation*

Proof.

1. **(reflectivity)** *If* $(B, G) \in \mathcal{A}$ *then*
    a. $B \subseteq B$
    b. $G \subseteq G$

c. $x \in B \wedge y \in B \backslash B = \emptyset \Rightarrow (x, y) \in G$ *is satisfied because* $y \in \emptyset$ *can not occur so* $x \in B \wedge y \in B \backslash B$ *is always false.*

proving $(B, G) \prec (B, G)$

2. **(anti-symmetry)** *If* $(B, G) \prec (B', G') \wedge (B', G') \prec (B, G) \Rightarrow B \subseteq B' \subseteq B \wedge G \subseteq G' \subseteq G \Rightarrow B = B' \wedge G = G' \Rightarrow (B, G) = (B', G')$ *proving anti-symmetry.*

3. **(transitivity)** *If* $(B, G) \prec (B', G') \wedge (B', G') \prec (B'', G'')$ *then we have*

   a. $B \subseteq B' \wedge B' \subseteq B'' \Rightarrow B \subseteq B''$

   b. $G \subseteq G' \wedge G' \subseteq G'' \Rightarrow G \subseteq G''$

   c. *If* $x \in B \wedge y \in B'' \backslash B$ *we have then for* y *the following cases*

      i. $y \in B' \Rightarrow y \in B' \backslash B \Rightarrow (x, y) \in G' \underset{G' \subseteq G''}{\Rightarrow} (x, y) \in G''$

      ii. $y \notin B' \Rightarrow y \in B'' \backslash B' \Rightarrow (x, y) \in G''$

   So in all cases we have $(x, y) \in G''$

proving $(B, G) \prec (B'', G'')$ □

We now have the following lemma.

Lemma 2.221. *If* $\mathcal{C} \subseteq \mathcal{A}$ *is a chain in* $\langle \mathcal{A}, \prec \rangle$ *then if*

$$\begin{aligned} B_{\mathcal{C}} &= \bigcup_{(B,G) \in \mathcal{C}} B \\ G_{\mathcal{C}} &= \bigcup_{(B,G) \in \mathcal{C}} G \end{aligned}$$

we have $(B_{\mathcal{C}}, G_{\mathcal{C}}) \in \mathcal{A}$

Proof.

1. $\forall (B, G) \in \mathcal{C}$ we have $B \subseteq A$ so if $x \in B_{\mathcal{C}}$ then $\exists (B, G) \in \mathcal{C} \vdash x \in B \subseteq A \Rightarrow x \in A$ and thus $B_{\mathcal{C}} \subseteq A$

2. We must prove now that $G_{\mathcal{C}}$ is a order relation on $B_{\mathcal{C}}$

   a. **(reflectivity)** If $x \in B_{\mathcal{C}} \Rightarrow \exists (B, G) \in \mathcal{C} \vdash x \in B \Rightarrow (x, x) \in G \Rightarrow (x, x) \in G_{\mathcal{C}}$

   b. **(anti-symmetry)** If $(x, y) \in G_{\mathcal{C}} \wedge (y, x) \in G_{\mathcal{C}}$ then $\exists (B, G), (B', G') \in \mathcal{C} \vdash (x, y) \in G \wedge (y, x) \in G'$ now because $\mathcal{C}$ is a chain we have either

      i. $(B, G) \prec (B', G')$ then $G \subseteq G' \Rightarrow (x, y), (y, x) \in G' \Rightarrow x = y$

      ii. $(B', G') \prec (B, G)$ then $G' \subseteq G \Rightarrow (x, y), (y, x) \in G \Rightarrow x = y$

   c. **(transitivity)** If $(x, y), (y, z) \in G_{\mathcal{C}}$ then $\exists (B, G), (B', G') \in \mathcal{C} \vdash (x, y) \in G \wedge (y, z) \in G'$. Now because $\mathcal{C}$ is a chain we have either

      i. $(B, G) \prec (B', G') \Rightarrow G \subseteq G' \Rightarrow (x, y), (y, z) \in G' \Rightarrow (x, z) \in G' \Rightarrow (x, z) \in G_{\mathcal{C}}$

ii. $(B', G') \prec (B, G) \Rightarrow G' \subseteq G \Rightarrow (x, y), (y, z) \in G \Rightarrow (x, z) \in G \Rightarrow (x, z) \in G_{\mathcal{C}}$

3. Now we must prove well-ordering. Suppose that $D \subseteq B_{\mathcal{C}}$ and $D \neq \emptyset$ then there exists a $x \in D \Rightarrow x \in B_{\mathcal{C}} \Rightarrow \exists (B, G) \in \mathcal{C} \vdash x \in B \Rightarrow D \bigcap B \neq \emptyset$. Now $D \bigcap B \subseteq B$ hence by well-ordering of $\langle B, G \rangle$ we have that there exists a least element $b \in B$ in $B \Rightarrow \forall y \in B \vDash (b, y) \in G \Rightarrow \forall y \in B \bigcap D \vDash (b, y) \in G$. We prove now that b is the least element of D.

   Proof. If $x \in D \Rightarrow \exists (B', G')$ with $x \in B'$ then we have the following cases

   a. $x \in B \Rightarrow x \in B \bigcap D \Rightarrow (b, x) \in G \Rightarrow (b, x) \in G_{\mathcal{C}}$

   b. $x \notin B \Rightarrow x \in B' \backslash B \wedge b \in B$ now we have two cases

      i. $(B, G) \prec (B', G') \underset{b \in B \wedge x \in B' \backslash B}{\Rightarrow} (b, x) \in B' \Rightarrow (b, x) \in G_{\mathcal{C}}$

      ii. $(B', G') \prec (B, G) \Rightarrow B' \subseteq B \Rightarrow x \in B$ contradicting $x \notin B$ so this case does not apply.

   So in all cases that are valid we have $(b, x) \in G_{\mathcal{C}}$ proving that b is the least element of D and thus proving that $\langle B_{\mathcal{C}}, G_{\mathcal{C}} \rangle$ is a well-ordered set. □ □

Next we have the finally lemma before the Well-Ordering theorem

Lemma 2.222. *If $\mathcal{C}$ is a chain in $\langle \mathcal{A}, \prec \rangle$ then $(B_{\mathcal{C}}, G_{\mathcal{C}})$ is a upper bound of $\mathcal{C}$*

Proof. If $(B, G) \in \mathcal{C}$ then clearly $B \subseteq B_{\mathcal{C}}$ and $G \subseteq G_{\mathcal{C}}$ so to prove that $(B, G) \prec (B_{\mathcal{C}}, G_{\mathcal{C}})$ we must prove that if $x \in B$ and $y \in B_{\mathcal{C}} \backslash B$ that $(x, y) \in G_{\mathcal{C}}$. So assume that $x \in B$ and $y \in B_{\mathcal{C}} \backslash B$ then $\exists (B', G') \vdash y \in B'$. We have then as $y \notin B$ that $B' \nsubseteq B$ hence as $\mathcal{C}$ is a chain and we can't have $(B', G') \prec (B, G)$ we must have $(B, G) \prec (B', G')$ and thus as $B \subseteq B'$ we have $y \in B' \backslash B \Rightarrow (x, y) \in G' \Rightarrow (x, y) \in B_{\mathcal{C}}$ proving that $(B, G) \prec (B_{\mathcal{C}}, G_{\mathcal{C}})$. □

Finally we can prove our well-ordering theorem. Using Zorn's lemma (2.216) we have a maximal element $(B_m, G_m) \in \mathcal{A}$. We prove now that $B_m = A$ proving that $(A, G_m) \in \mathcal{A}$ and that $\langle A, G_m \rangle$ is well-ordered.

Proof. Assume that $B_m \neq A$ then as $B_m \subseteq A$ we have $\exists x \in A \backslash B_m$ define then $G^* = G_m \bigcup \{(b, x) | b \in B_m\} \bigcup \{(x, x)\}$ (a disjoint union), we prove then that $\langle B_m \bigcup \{x\}, G^* \rangle$ is a well-ordered set

Proof. Note that if $(x, r) \in G^*$ then we have the following possibilities

1. $(x, r) \in G_m$ which is impossible because $G_m \subseteq B_m \times B_m$ and $x \notin B_m$
2. $(x, r) \in \{(b, x) | b \in B_m\}$ which is impossible as $x \notin B_m$
3. $(x, r) \in \{(x, x)\} \Rightarrow r = x$

So from $(x, r) \in G^*$ it follows that $(x, r) = (x, x) \in \{(x, x)\}$ we now prove that G^* is a well-ordering.

1. **(reflexive)** If $y \in B_m \bigcup \{x\}$ then we have the following cases

   a. $y \in B_m \Rightarrow (y, y) \in G_m \Rightarrow (y, y) \in G^*$

b. $y\in\{x\}\Rightarrow(y,y)=(x,x)\in G^{*}$

2. **(anti-symmetry)** If $(r,s),(s,r)\in G^{*}$ we have either

   a. $(r,s)\in G_m \underset{G_m\subseteq B_m\times B_m}{\Rightarrow} r,s\in B_m \underset{r,s\neq x\wedge(s,r)\in G^{*}}{\Rightarrow}(s,r)\in G_m\Rightarrow r=s$

   b. $(r,s)\in\{(b,x)|b\in B_m\}\Rightarrow s=x \underset{(x,r)\in G^{*}}{\Rightarrow} r=x\Rightarrow r=s$

   c. $(r,s)\in\{(x,x)\}\Rightarrow r=x=s$

3. **(transitive)** If $(r,s),(s,t)\in G^{*}$ then we have the following cases

   a. $(r,s)\in G_m$ we have then the following sub cases

      i. $(s,t)\in G_m\Rightarrow(r,t)\in G_m\Rightarrow(r,t)\in G^{*}$

      ii. $(s,t)=(b,x)\Rightarrow t=x\Rightarrow(r,t)=(r,x)\in G^{*}\Rightarrow(r,t)\in G^{*}$

      iii. $(s,t)=(x,x)\Rightarrow t=x\Rightarrow(r,t)=(r,x)\in G^{*}\Rightarrow(r,t)\in G^{*}$

   b. $(r,s)\in\{(b,x)|b\in B_m\}\Rightarrow s=x\Rightarrow(x,t)\in G^{*}\Rightarrow t=x\Rightarrow(r,t)\in\{(b,x)|b\in B_m\}\Rightarrow(r,t)\in G^{*}$

   c. $(r,s)\in\{(x,x)\}\Rightarrow r=s=x\Rightarrow(x,t)\in G^{*}\Rightarrow t=x\Rightarrow(r,t)=(x,x)\in G^{*}$

so G^{*} is indeed a order relation. Now if $C\subseteq B_m\bigcup\{x\}$ is a nonempty set then we have the following possibilities

1. $C\bigcap B_m\neq\emptyset\Rightarrow\emptyset\neq C\bigcap B_m\subseteq B_m \underset{G_m\,\text{is well-ordered}}{\Rightarrow} \exists l\in C\bigcap B_m\vdash\forall y\in C\bigcap B_m\vDash(l,y)\in G_m$. Now if $y\in C$ then we have the following possibilities

   a. $y\in B_m\Rightarrow y\in C\bigcap B_m\Rightarrow(l,y)\in G_m\Rightarrow(l,y)\in G^{*}$

   b. $y\notin B_m\Rightarrow y=x\Rightarrow(l,y)\in\{(b,x)|b\in B_m\}\Rightarrow(l,y)\in G^{*}$

   So we conclude that C has a least element.

2. $C\bigcap B_m=\emptyset\Rightarrow C=\{x\}\Rightarrow x$ is the least element of C.

So we have that $(B_m\bigcup\{x\},G^{*})\in\mathcal{A}$

Next we have $(B_m,G_m)\prec(B_m\bigcup\{x\},G^{*})$ as

1. $B_m\subseteq B_m\bigcup\{x\}$
2. $G_m\subseteq G^{*}$
3. If $r\in B_m$ and $s\in(B_m\bigcup\{x\})\backslash B_m\Rightarrow s=x\Rightarrow(r,s)\in\{(b,x)|b\in B_m\}\Rightarrow(r,s)\in G^{*}$

But as $x\notin B_m$ we have $B_m\neq B_m\bigcup\{x\}$ and thus $(B_m,G_m)\neq(B_m\bigcup\{x\},G^{*})$ so we have a contradiction as by maximality of (B_m,G_m) in $\langle\mathcal{A},\prec\rangle$ we would have $(B_m,G_m)=(B_m\bigcup\{x\},G^{*})$. So we must conclude that $B_m=A$ □

So finally $(A,G_m)\in\mathcal{A}\Rightarrow\langle A,G_m\rangle$ is a well-ordered set. □

□

Using the above we have proved that

Axiom Of Choice $\Rightarrow$ Hausdorff's Maximal Principle $\Rightarrow$ Zorn's Lemma $\Rightarrow$ Well-Ordering Theorem.

We prove now that the Axiom of Choice follows from the Well-Ordering Theorem.

Theorem 2.223. *The Well-Ordering Theorem implies the Axiom of Choice*

Proof. If A is any set then choose a well-ordering $\langle A, \leqslant \rangle$ of A. Define then c: $\mathcal{P}'(A) \rightarrow A$ by $c = \{(B, b) | B \in \mathcal{P}'(A) \wedge b \text{ is least element of } B\} \subseteq \mathcal{P}'(A) \times A$. We have then that if $(B, b), (B, b') \in c \underset{\text{least element is unique}}{\Rightarrow} b = b'$ proving that c is a partial function. And if $B \in \mathcal{P}'(A) \Rightarrow \emptyset \neq B \subseteq A \underset{\text{well}-\text{ordering}}{\Rightarrow} \exists b \vdash b \text{ is least element in } B \Rightarrow (B, b) \in c$ proving that $\text{dom}(c) = \mathcal{P}'(A)$ and thus that $c: \mathcal{P}'(A) \rightarrow A$ is a function. Finally if $B \in \mathcal{P}'(A)$ we have $c(B) \in B$ as the least element of a set is in the set itself. $\square$

So we have proved the following equivalences.

Theorem 2.224. *The following statements are equivalent*

1. *Axiom of Choice*
2. *Hausdorff's Maximal Principle*
3. *Zorn's Lemma*
4. *Every set can be well-ordered*

For the rest of this document we assume that the Axiom of Choice is valid and thus also that the Hausdorff's Maximal Principle, Zorn's Lemma and 'Every set can be well-ordered' holds true.

As an example of applying the Axiom of Choice let's prove the following theorem.

Theorem 2.225. *Assume X, Y sets and $f: X \rightarrow Y$ a function then there exists a $Z \subseteq X$ such that $f_{|Z}: X \rightarrow Y$ is injective and $f_{|Z}(X) = f(X)$ (which ensures that $f_{|Z}$: $Z \rightarrow f(X)$ is a bijection).*

Proof. First define $\mathcal{A} = \{f^{-1}(\{y\}) | y \in f(X)\}$. If $A \in \mathcal{A}$ there exists a $y \in f(X)$ such that $A = f^{-1}(\{y\}) \Rightarrow A \subseteq X$, as $y \in f(X)$ there exists a $x \in X$ such that $y = f(x) \Rightarrow x \in f^{-1}(\{y\}) = A \Rightarrow A \neq \emptyset \Rightarrow A \in \mathcal{P}'(X)$ so we conclude that $\mathcal{A} \subseteq \mathcal{P}'(X)$. By the Axiom of Choice (see 2.198) there exists a choice function $c: \mathcal{P}'(X) \rightarrow X$ such that if $A \in \mathcal{P}'(X)$ then $c(A) \in A$. Take now $Z = c(\mathcal{A})$ and consider $f_{|Z}: Z \rightarrow Y$ and prove that is injective. Let $z_1, z_2 \in Z$ is such that $f_{|Z}(z_1) = f_{|Z}(z_2) \Rightarrow f(z_1) = f(z_2)$. As $z_1, z_2 \in Z = c(\mathcal{A})$ there exists $A_1, A_2 \in \mathcal{A}$ such that $z_1 = c(A_1) \in A_1, z_2 = c(A_2) \in A_2$ and thus as $A_1, A_2 \in \mathcal{A}$ there exists $y_1, y_2 \in f(X)$ such that $A_1 = f^{-1}(\{y_1\}), A_2 = f^{-1}(\{y_2\})$ so $z_1 \in f^{-1}(\{y_1\}), z_2 \in f^{-1}(\{y_2\}) \Rightarrow f(z_1) = y_1, f(z_2) = y_2 \underset{f(z_1) = f(z_2)}{\Rightarrow} y_1 = y_2 \Rightarrow A_1 = f^{-1}(\{y_1\}) = f^{-1}(\{y_2\}) = A_2 \Rightarrow z_1 = c(A_1) = c(A_2) = z_2 \Rightarrow z_1 = z_2$ proving that $f_{|Z}: Z \rightarrow Y$ is injective.

If now $y \in f(X)$ then $f^{-1}(\{y\}) \in \mathcal{A} \Rightarrow x = c(f^{-1}(\{y\})) \in c(\mathcal{A}) = Z \Rightarrow x \in Z$ and as also $x = c(f^{-1}(\{y\})) \in f^{-1}(\{y\})$ we have $f(x) = y$ so we have fount a $x \in Z$ such that $y = f(x) \Rightarrow f(X) \subseteq f(Z)$. From $Z \subseteq X$ we have $f(Z) \subseteq f(X)$ and we conclude thus that $f(X) = f(Z)$ as must be proved. □

Finally we prove that Zorn's lemma is also valid if we replace partial ordered sets by pre-ordered set.

Theorem 2.226. *Let $\langle A, \leqslant \rangle$ be a pre-ordered set (see 2.131) such that every chain $C \subseteq A$ has a upper bound (see 2.165) then A has a maximal element (see 2.160)*

Proof. Use 2.138 to construct the partial ordered set $\langle A/\sim, \leqslant \rangle$ where $x \sim y$ iff $x \leqslant y \wedge y \leqslant x$ and where $\sim[x] < \sim[y]$ iff $x < y$. Let now $C \subseteq A/\sim$ be a chain in $A/\sim$ (in the pre-order sense) . Form then $C' = \bigcup_{S \in C} S \subseteq A$. If $x, y \in C'$ then there exists $\sim[x'], \sim[y'] \in C$ such that $x \in \sim[x'], y \in \sim[y']$ or $x \leqslant x' \wedge x' \leqslant x$ and $y \leqslant y' \wedge y' \leqslant y$. As C is a chain we have either:

1. $\sim[x'] \leqslant \sim[y'] \underset{2.138}{\Rightarrow} x' \leqslant y' \Rightarrow x \leqslant y$
2. $\sim[y'] \leqslant \sim[x'] \underset{2.138}{\Rightarrow} y' \leqslant x' \Rightarrow y \leqslant x$

proving that C' is a chain in A. By the hypothesis there exists a upper bound u of C' in A, in other words $\exists u \in U$ such that $\forall a \in A$ we have $a \leqslant u$. Consider now $\sim[u]$ and take $\sim[z] \in C$ then as $z \in \sim[z]$ we have $z \in C'$ and thus $z \leqslant u$ proving that $\sim[z] \leqslant \sim[u]$. This proves that every chain in $A/\sim$ has a upper bound and thus by Zorn's lemma there exists a maximal element $\sim[m] \in A$ of A. If now $x \in A$ with $m \leqslant x$ then $\sim[m] \leqslant \sim[x]$ so by maximality of $\sim[m]$ we have then that $\sim[m] = \sim[x]$ giving $x \leqslant m$ and thus by 2.160 that m is a maximal element in the pre-order of A. □

Chapter 3

Algebraic Constructs

3.1 Groups

Definition 3.1. *A semi-group is a pair* $\langle G, \odot \rangle$ *where* G *is a set and* $\odot: G \times G \rightarrow G$ *a function such that (here* $\odot(x, y)$ *is noted as* $x \odot y$*)*

1. **(Neutral Element)** $\exists e \in G \vdash \forall x \in G \vdash x \odot e = e \odot x$
2. **(Associativity)** $\forall x, y, z \in G$ *we have* $x \odot (y \odot z) = (x \odot y) \odot z$

Lemma 3.2. *If* $\langle G, \odot \rangle$ *is a semi-group then* $G \neq \emptyset$

Proof. This follows from (1) in the definition. □

Theorem 3.3. *If* $\langle G, \odot \rangle$ *is a semi-group then there exists only one neutral element.*

Proof. If e, e' are neutral elements then $e = e' \odot e = e'$ □

Definition 3.4. *A group is a semi-group* $\langle G, \odot \rangle$ *with the extra condition*

1. **(Inverse Element)** $\forall x \in G \vdash \exists y \in G \vDash x \odot y = e = y \odot x$

Theorem 3.5. *If* $\langle G, \odot \rangle$ *is a group then* $\forall x \in G \vdash \exists! y \in G \Rightarrow x \circ y = e = y \odot x$ *this unique element is noted as* x^{-1}

Proof. Take $x \in G$ and suppose that y, y' is such that $x \odot y = e = y \odot x$ and $x \odot y' = e = y' \odot x$ then $y = y \odot e = y \odot (x \odot y') = (y \odot x) \odot y' = e \odot y' = y'$ □

Theorem 3.6. *If* $\langle G, \odot \rangle$ *is a group then* $\forall x \in G$ *we have* $(x^{-1})^{-1} = x$

Proof. If $x \in G$ then $x \odot x^{-1} = e = x^{-1} \odot x$ and $(x^{-1})^{-1} \odot x^{-1} = e = x^{-1} \odot (x^{-1})^{-1}$, using the previous theorem we have then $x = (x^{-1})^{-1}$ □

Definition 3.7. *A group or semi-group* $\langle G, \odot \rangle$ *is* ***abelian*** *or* ***commutative****.If the following is satisfied*

1. **(Commutativity)** $\forall x, y \in G \vdash x \odot y = y \odot x$

Definition 3.8. *If* $\langle G, \odot \rangle$ *is a semi-group then* $F \subseteq G$ *is a sub-semi-group of* G *if*

1. $\forall x, y \in F \vDash x \odot y \in F$

2. $e \in F$

Theorem 3.9. *If* $\langle G, \odot \rangle$ *is a semi-group then if* $F \subseteq G$ *is a sub-semi-group then* $\langle F, \odot_{|F \times F} \rangle$ *is a semi-group. For simplicity we note this as* $\langle F, \odot \rangle$

Proof.

1. If $x, y \in F \Rightarrow x \odot y \in F \Rightarrow x \odot_{|F \times F} y \in F \Rightarrow \odot_{|F \times F}: F \times F \rightarrow F$ is a function
2. $e \in F \Rightarrow x \odot_{|F \times F} e = x \odot e = x = e \odot x = e \odot_{|F \times F} x$
3. If $x, y, z \in F \Rightarrow x \odot_{|F \times F} (y \odot_{|F \times F} z) = x \odot (y \odot z) = (x \odot y) \odot z = (x \odot_{|X \times Y} y) \odot_{|F \times F} z$ □

Definition 3.10. *If* $\langle G, \odot \rangle$ *is a group then* $F \subseteq G$ *is a sub-group of* G *if*

1. $\forall x, y \in F \vDash x \odot y \in F$

2. $\forall x \in G \vDash x^{-1} \in F$

3. $e \in F$ *where* e *is the unit in* G*.*

Theorem 3.11. *If* $\langle G, \odot \rangle$ *is a group then if* $F \subseteq G$ *is a sub-group we have that* $\langle F, \odot_{|F \times F} \rangle$ *is a group noted as* $\langle F, \odot \rangle$ *for simplicity.*

Proof. As $e \in F$ we have using 3.9 that $\langle F, \odot \rangle$ is a semi-group. Also if $x \in F \Rightarrow x^{-1} \in F = x \odot_{|F \times F} x^{-1} = x \odot x^{-1} = e = x^{-1} \odot x = x^{-1} \odot_{|F \times F} x$ □

Definition 3.12. *If* $\langle F, \odot \rangle$ *and* $\langle G, \oplus \rangle$ *are semi-groups then a* ***group homeomorphism*** *is a function* $f: F \rightarrow G$ *such that* $\forall x, y \in F$ *we have* $f(x \odot y) = f(x) \oplus f(y)$

Theorem 3.13. *If* $\langle F, \odot \rangle$ *and* $\langle G, \oplus \rangle$ *are groups with units* e_F, e_G *and* $f: F \rightarrow G$ *a group homeomorphism then we have*

1. $f(e_F) = f(e_G)$

2. $\forall a \in F$ *we have* $f(a^{-1}) = f(a)^{-1}$

Proof.

1. $f(e_F) = f(e_F \cdot e_F) = f(e_F) \cdot f(e_F) \underset{\text{multiply both sides by } f(e_F)^{-1}}{\Rightarrow} e_G = f(e_F)$
2. If $a \in F \Rightarrow f(a^{-1} \cdot a) = f(e_F) \underset{(1)}{=} e_G \underset{(2)}{=} f(e_F) = f(a \cdot a^{-1}) \Rightarrow f(a^{-1}) \cdot f(a) = e_G = f(a) \cdot f(a^{-1}) \Rightarrow f(a^{-1}) = f(a)^{-1}$ □

Definition 3.14. *If* $\langle F, \odot \rangle$ *and* $\langle G, \oplus \rangle$ *are semi-groups then a* ***group isomorphism*** *is a function* $f: F \rightarrow G$ *that is a group homeomorphism and a bijection.*

Theorem 3.15. *If* $\{\langle F_i, \odot_i \rangle\}_{i \in I}$ *is a family of (abelian) groups (or semi-groups) then we have:*

1. $\forall x, y \in \prod_{i \in I} F_i$ *we have:*

  a. $\forall i \in I$ *that* $x(i) \odot_i y(i) \in F_i$

b. *$x \odot y: I \to \bigcup_{i \in I} F_i$ defined by $i \to x(i) \odot_i y(i)$ is a element of $\prod_{i \in I} F_i$*

2. *$\odot: \prod_{i \in I} F_i \times \prod_{i \in I} F_i \to \prod_{i \in I} F_i$ defined by $(x, y) \to x \odot y$ is a function [here we use $x \odot y$ to actually note the function $\langle x \odot y, I, \bigcup_{i \in I} F_i \rangle$]*

3. *$\langle \prod_{i \in I} F_i, \odot \rangle$ is a (abelian) group (or semi group) with neutral element $0: I \to \bigcup_{i \in I} F_i$ defined by $i \to 0_i \in F_i$ (where 0_i is the neutral element in F_i) and if $x \in \prod_{i \in I} F_i$ then $-x: I \to \bigcup_{i \in I} F_i$ defined by $i \to -x(i) \in F_i$ (where $-x(i)$ is the inverse element of $x(i)$)*

Proof.

1. If $x, y \in \prod_{i \in I} F_i$ then we have

   a. As $x(i), y(i) \in F_i$ by the definition of $\prod_{i \in I} F_i$ we have as $\odot: F_i \odot_i F_i \to F_i$ is a function we have that $x(i) \odot_i y(i) \in F_i$

   b. As $\forall i \in I$ we have by $x(i) \odot_i y(i) \in F_i$ we have by definition of $\prod_{i \in I} F_i$ that $x \odot y \in \prod_{i \in I} F_i$

2. This follows from the definition of $\odot$ and the fact that $x \odot y \in F_i$

3. We have if $\{\langle F_i, \odot_i \rangle\}_{i \in I}$ is a family of semi-groups then we have

   - **(associativity)** $\forall i \in I$ we have $((x \odot y) \odot z)(i) = (x \odot y)(i) \odot_i z(i) = (x(i) \odot_i y(i)) \odot_i z(i) \underset{\langle F_i, \odot_i \rangle \text{ is a semi-group}}{=} x(i) \odot_i (y(i) \odot_i z(i)) = x(i) \odot_i (y \odot z)(i) = (x \odot (y \odot z))(i)$ resulting in $(x \odot y) \odot z = x \odot (y \odot z)$

   - **(neutral element)** $\forall i \in I$ we have $(x \odot 0)(i) = x(i) \odot_i 0(i) = x(i) \odot_i 0_i = x(i) = 0_i \odot_i x(i) = (0 \odot x)$ proving that $0 \odot x = x = x \odot 0$

   If $\{\langle F_i, \odot_i \rangle\}_{i \in I}$ is also a group then we have

   - **(inverse element)** $\forall i \in I$ we have $(x \odot (-x))(i) = x(i) \odot_i ((-x)(i)) = x(i) \odot (-(x(i))) = 0_i = 0(i) = 0_i = (-x(i)) \odot x(i) = (-x)(i) \circ_i x(i) = ((-x) \odot x)(i)$ giving $x \circ (-x) = 0 = (-x) \odot x$

   and in case of a abelian family we have

   - **(commutativity)** $\forall i \in I$ we have $(x \odot y)(i) = x(i) \odot_i y(i) = y(i) \odot_i x(i) = (y \odot x)(i)$ giving $x \odot y = y \odot x$ $\square$

Definition 3.16. *Let $\langle G, \odot \rangle$ be a group with neutral element 0_G and let X be a set then we have the following definitions:*

1. *A* ***left group action*** *is a function $\triangleright: G \times X \to X$ where $\triangleright(g, x) \underset{\text{noted}}{=} g \triangleright x$ such that*

   a. *$\forall x \in X$ we have $0_G \triangleright x = x$*

   b. *$\forall g, g' \in G$ and $\forall x \in X$ we have $(g \odot g') \triangleright x = g \triangleright (g' \triangleright x)$*

2. *A* ***right group action*** *is a function $\triangleleft: X \times G \to X$ where $\triangleleft(x, g) \underset{\text{noted}}{=} x \triangleleft g$ such that*

   a. *$\forall x \in X$ we have $x \triangleleft 0_G = x$*

b. $\forall g, g' \in G$ *and* $\forall x \in X$ *we have* $x \lhd (g \odot g') = (x \lhd g) \lhd g'$

Note 3.17. If $\langle G, \odot \rangle$ is a group and X as set with a left group action $\rhd$ then you think that we can always define a right group action $\lhd$ as follows: $\lhd: X \times G \rightarrow X$ is defined by $(x, g) \rightarrow x \lhd g \equiv g \rhd g$, however this will in general not give you a right group action as we have not in general that 2 (b) is not fulfilled for $x \lhd (g \odot g') \underset{\text{definition}}{=} (g \odot g') \rhd x = g \rhd (g' \rhd x) \underset{\text{definition}}{=} (g' \rhd x) \lhd g \underset{\text{definition}}{=} (x \lhd g') \lhd g$ which in general is not the same as $(x \lhd g) \lhd g'$.

Definition 3.18. *Let* $\langle G, \odot \rangle$ *be a group,* X *a set then if* $\rhd$ *(*$\lhd$*) is a left (right) group action then we define if* $g \in G$

1. $g_{\rhd}: X \rightarrow X$ *by* $x \rightarrow g_{\rhd}(x) = g \rhd x$
2. $g_{\lhd}: X \rightarrow X$ *by* $x \rightarrow g_{\lhd}(x) = x \lhd g$

Definition 3.19. *Let* $\langle G, \odot \rangle$ *be a group with neutral element* 0_G *and let* X *be a set then we have the following definitions for a left (right) group action* $\rhd$ $(\lhd)$*:*

1. $\rhd$ *(or* $\lhd$*) is* ***faithful*** *iff* $g_{\rhd} = 1_X$ *(or* $g_{\lhd} = 1_X$*) if and only if* $g = 0_G$*where* 1_X: $X \rightarrow X$ *is the identity mapping. Equivalently this means that* $\{g \in G | \forall x \in X | g \rhd x = x\} = \{0_G\}$ *(or* $\{g \in G | \forall x \in X | x \lhd g = x\} = \{0_G\}$*)*
2. $\rhd$ *(or* $\lhd$*) is* ***transitive*** *iff* $\forall x_1, x_2$ *there exists a* $g \in G$ *such that* $g \rhd x_1 = x_2$ *(or* $x_1 \rhd g = x_2$*)*
3. $\rhd$ *(or* $\lhd$*) is* ***free*** *iff* $\forall x \in X$ *we have* $\{g \in G | g \rhd x = x\} = \{0_G\}$ *(or* $\{g \in | x \lhd g\} = \{0_G\}$*)*

3.2 Rings

Definition 3.20. (Ring) *A triple* $\langle R, \oplus, \odot \rangle$ *is a ring iff*

1. R *is a set*
2. $\oplus: R \times R \rightarrow R$ *is a function*
3. $\odot: R \times R \rightarrow R$ *is a function*
4. $\langle R, \oplus \rangle$ *is a abelian group*
   a. **(Associative)** $\forall a, b, c \in R \vDash a \oplus (b \oplus c) = (a \oplus b) \oplus c$
   b. **(Neutral element)** $\exists 0 \in R \vdash \forall a \in R \vDash a \oplus 0 = a = 0 \oplus a$
   c. **(Inverse)** $\forall a \in R$ *there* $\exists -a \in R \vDash a \oplus (-a) = 0 = (-a) \oplus a$
   d. **(Commutative)** $\forall a, b \in R \vdash a \oplus b = b \oplus a$
5. **(Distributive)** $\forall a, b, c \in R \vDash a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c)$
6. **(Neutral element)** $\exists 1 \in R \vdash \forall a \in R \vDash a \odot 1 = a = 1 \odot a$
7. **(Commutative)** $\forall a, b \in R \vDash a \odot b = b \odot a$
8. **(Associative)** $\forall a, b, c \in R \vDash a \odot (b \odot c) = (a \odot b) \odot c$

Note that in a ring we have also that $\langle R, \odot \rangle$ is a abelian semi-group.

Definition 3.21. *If $\langle R, \oplus, \odot \rangle$ is a ring then a* ***zero divisor*** *is a $x \in R \backslash \{0\}$ so that there exists a $b \in R \backslash \{0\}$ such that $a \cdot b = 0$.*

Definition 3.22. *A ring $\langle R, \oplus, \odot \rangle$ is a* ***integral domain*** *if it does not contain a* ***zero divisor***

Definition 3.23. (Sub-Ring) *If $\langle R, \oplus, \odot \rangle$ is a ring then a subset $S \subseteq R$ is a sub ring iff*

1. *$\forall a, b \in S$ we have $a \oplus b \in S$*
2. *$\forall a, b \in S$ we have $a \odot b \in S$*
3. *$\forall a \in S$ we have $-a \in S$*
4. *$1 \in S$ where 1 is the unit of multiplication $\odot$*
5. *$0 \in S$ where 0 is the unit of multiplication $\oplus$*

Theorem 3.24. *If $\langle R, \oplus, \odot \rangle$ is a ring and $S \subseteq R$ a sub-ring then $\langle S, \oplus_{|S \times S}, \odot_{|S \times S} \rangle$ is ring. For simplicity we note this ring as $\langle S, \oplus, \odot \rangle$.*

Proof.

1. As $S \subseteq R$ then by 1.64 S is a set.
2. As $\oplus: R \times R \rightarrow R$ is a function we have by 2.27 that $\oplus_{|S \times S}: S \times S \rightarrow S$ is a function.
3. As $\odot: R \times R \rightarrow R$ is a function we have by 2.27 that $\odot_{|S \times S}: S \times S \rightarrow S$ is a function.
4. $\langle S, \oplus_{|S \times S} \rangle$ is a abelian group
   a. **(Associative)** $\forall a, b, c \in S$ we have $(a \oplus_{|S \times S} b) \oplus_{|S \times S} c = (a \oplus b) \oplus c = a \oplus (b \oplus c) = a \oplus_{|S \times S} (b \oplus_{|S \times S} c)$
   b. **(Neutral element)** As $0 \in S$ we have $\forall a \in S \Rightarrow a \oplus_{|S \times S} 0 = a \oplus 0 = a = 0 \oplus a = 0 \oplus_{|S \times S} a$
   c. **(Inverse element)** If $a \in S$ then there exists a $-a \in S$ with $a \oplus (-a) = 0 = (-a) \oplus a \Rightarrow a \oplus_{|S \times S} (-a) = a \oplus (-a) = 0 = (-a) \oplus a = (-a) \oplus_{|S \times S} a$
   d. **(Commutativity)** If $a, b \in S$ then $a \oplus_{|S \times S} b = a \oplus b = b \oplus a = b \oplus_{|S \times S} a$
5. **(Distributivity)** $\forall a, b, c \in S$ we have $a \odot_{|S \times S} (b \oplus_{|S \times S} c) = a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c) = (a \odot_{|S \times S} b) \oplus_{|S \times S} (a \odot_{|S \times S} c)$
6. **(Neutral Element)** As $1 \in S$ then we have if $a \in S$ that $1 \odot_{|S \times S} a = 1 \odot a = a \odot 1 = a \odot_{|S \times S} 1$
7. **(Commutativity)** If $a, b \in S$ then $a \odot_{|S \times S} b = a \odot b = b \odot a = b \odot_{|S \times S} a$
8. **(Associative)** If $a, b, c \in S$ then $a \odot_{|S \times S} (b \odot_{|S \times S} c) = a \odot (b \odot c) = (a \odot b) \odot c = (a \odot_{|S \times S} b) \odot_{|S \times S} c$ □

Theorem 3.25. *If $\langle R, \oplus, \odot \rangle$ is a ring with neutral element 0 in $\langle R, + \rangle$ then $\forall x \in R$ we have $x \odot 0 = 0 \odot x = 0$*

Proof. If $x \in R$ then $0 = (0 \odot x) \oplus (-(0 \odot x)) \underset{0 \oplus 0 = 0}{=} ((0 \oplus 0) \odot x) \oplus (-(0 \odot x)) = [(0 \odot x) \oplus (0 \odot x)] \oplus (-(0 \odot x)) = 0 \odot x \oplus [(0 \odot x) + (-(0 \odot x))] = 0 \odot x + 0 = 0 \odot x$ □

Definition 3.26. *If $\langle A, \oplus_A, \odot_A \rangle$ and $\langle B, \oplus_B, \odot_B \rangle$ are rings then a function f: $A \rightarrow B$ is a ring homeomorphism iff*

1. $\forall a, b \in A$ *we have* $f(a \oplus_A b) = f(a) \oplus_B f(b)$
2. $\forall a, b \in A$ *we have* $f(a \odot_A b) = f(a) \odot_B f(b)$
3. $f(1_A) = 1_B$ *where* 1_A *is the multiplicative inverse in* A *and* 1_B *is the multiplicative inverse in* B.

Note that a ring homeomorphism $f: A \rightarrow B$ for the rings $\langle A, \oplus_A, \odot_A \rangle$, $\langle B, \oplus_B, \odot_B \rangle$ is automatically a group homeomorphism for the groups $\langle A, \oplus_A \rangle$, $\langle B, \oplus_B \rangle$. Using 3.13 we have then the following theorem

Theorem 3.27. *If $\langle A, \oplus_A, \odot_A \rangle$ and $\langle B, \oplus_B, \odot_B \rangle$ are rings with additive units 0_A, 0_B and $f: A \rightarrow B$ a ring homeomorphism then we have*

1. $f(0_A) = 0_B$
2. $\forall a \in A$ *we have* $f(-a) = -f(a)$

Definition 3.28. *If $\langle A, \oplus_A, \odot_A \rangle$ and $\langle B, \oplus_B, \odot_B \rangle$ are rings then a function f: $A \rightarrow B$ is a ring homeomorphism if it is a ring homeomorphism and a bijection.*

3.3 Fields

Definition 3.29. (Field) *A ring $\langle F, \oplus, \odot \rangle$ is a field if*

1. F *is a set*
2. $\oplus: F \times F \rightarrow F$ *is a function*
3. $\odot: F \times F \rightarrow F$ *is a function*
4. $\langle F, \oplus, \odot \rangle$ *is a Ring*
   a. $\langle F, \oplus \rangle$ *is a abelian group*
      i. **(Associative)** $\forall a, b, c \in F \vDash a \oplus (b \oplus c) = (a \oplus b) \oplus c$
      ii. **(Neutral element)** $\exists 0 \in F \vdash \forall a \in F \vDash a \oplus 0 = a = 0 \oplus a$
      iii. **(Inverse)** $\forall a \in F$ *there* $\exists -a \in F \vDash a \oplus (-a) = 0 = (-a) \oplus a$
      iv. **(Commutative)** $\forall a, b \in F \vdash a \oplus b = b \oplus a$
   b. **(Distributive)** $\forall a, b, c \in F \vDash a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c)$
   c. **(Neutral element)** $\exists 1 \in F \vdash \forall a \in F \vDash a \odot 1 = a = 1 \odot a$
   d. **(Commutative)** $\forall a, b \in F \vDash a \odot b = b \odot a$

e. **(Associative)** $\forall a, b, c \in F \vDash a \odot (b \odot c) = (a \odot b) \odot c$

5. **(Inverse for non zero element)** *If* $a \in F \backslash \{0\}$ *then* $\exists b \in F \vdash b \odot a = 1$

So a field is a ring that has also for every non-zero element a multiplicative inverse. We prove now that the inverse for non zero element is unique.

Theorem 3.30. *If* $\langle F, \oplus, \odot \rangle$ *is a field then* $\forall a \in F$ *with* $a \neq 0$ *there exists a* ***unique*** *inverse element, we note this inverse element by* a^{-1}.

Proof. Let $a \in F$ with $a \neq 0$ then by definition there exists a $b \in F$ such that $b \odot a = 1$, assume that there exists also a $b' \in F$ such that $b' \odot a = 1$. Then we have that $(b \odot a) \odot b' = b' \underset{\odot \text{ is associativity}}{\Rightarrow} b \odot (a \odot b') = b' \underset{\odot \text{ is commutative}}{\Rightarrow} b \odot (b' \odot a) = b' \Rightarrow b \odot 1 = b' \underset{\odot \text{ is commutative}}{\Rightarrow} 1 \odot b = b' \Rightarrow b = b'$ □

Definition 3.31. *If* $\langle F, \oplus, \odot \rangle$ *is a field then a subset* $S \subseteq F$ *is a sub-field iff the following is full filled*

1. $\forall a, b \in S \vDash a \oplus b \in S$
2. $\forall a, b \in S \vDash a \odot b \in S$
3. $\forall a \in S \vDash -a \in S$
4. $\forall a \in S \backslash \{0\} \vDash a^{-1} \in S$
5. $0 \in S$ *(where* 0 *is the additive neutral element of* F*)*
6. $1 \in S$ *(where* 1 *is the multiplicative neutral element of* F*)*

Theorem 3.32. *If* $\langle F, \oplus, \odot \rangle$ *is a field and* $S \subseteq F$ *a sub field then* $\langle S, \oplus_{|S}, \odot_{|S} \rangle$ *(for simplicity of notation we note this as* $\langle S, \oplus, \odot \rangle$*)*

Proof. Using 3.24 we have that $\langle S, \oplus_{|S}, \odot_{|S} \rangle$ is a sub ring so we only have to prove that every non-zero element in S has a multiplicative inverse. If $x \in S \backslash \{0\} \subseteq F \setminus \{0\} \Rightarrow x^{-1}$ exists, then by the definition of a sub-field $x^{-1} \in S$ and we have $x \odot_{|S} x^{-1} = x \odot x^{-1} = 1 = x^{-1} \odot x = x^{-1} \odot_{|S} x$. □

Definition 3.33. *If* $\langle F_1, \odot_1, \oplus \rangle$ *and* $\langle F_2, \odot_2, \oplus_2 \rangle$ *are fields with multiplicative units* $1_1, 1_2$ *then a function* $f: F_1 \rightarrow F_2$ *is a field homeomorphism iff*

1. $\forall a, b \in F_1$ *we have* $f(a \odot_1 b) = f(a) \odot_2 f(b)$
2. $\forall a, b \in F_1$ *we have* $f(a \oplus_1 b) = f(a) \oplus_2 f(b)$
3. $f(1_1) = 1_2$

If f *is also a bijection then we call* f *a field isomorphism.*

Note that a field homeomorphism $f: A \rightarrow B$ for the fields $\langle A, \oplus_A, \odot_A \rangle, \langle B, \oplus_B, \odot_B \rangle$ is automatically a group homeomorphism for the groups $\langle A, \oplus_A \rangle, \langle B, \oplus_B \rangle$. Using 3.13 we have then the following theorems

Theorem 3.34. *If* $\langle F_1, \odot_1, \oplus \rangle$ *and* $\langle F_2, \odot_2, \oplus_2 \rangle$ *are fields with multiplicative units* $1_1, 1_2$ *and* $f: F_1 \rightarrow F_2$ *is a field isomorphism then* $f^{-1}: F_2 \rightarrow F_1$ *is a field isomorphism*

Proof.

1. If $a, b \in F_2$ then $f((f^{-1})(a) \odot_1 (f^{-1})(b)) = f((f^{-1})(a)) \odot_2 f((f^{-1})(b)) = f((f^{-1})(a)) \odot_2 f((f^{-1})(b)) = a \odot_2 b$ hence $(f^{-1})(a \odot_2 b) = (f^{-1})(f((f^{-1})(a) \odot_1 (f^{-1})(b))) = (f^{-1})(a) \odot_1 (f^{-1})(b)$
2. If $a, b \in F_1$ then $f((f^{-1})(a) \oplus_1 (f^{-1})(b)) = f((f^{-1})(a)) \oplus_2 f((f^{-1})(b)) = a \oplus_2 b$ hence $(f^{-1})(a \oplus_2 b) = (f^{-1})(f((f^{-1})(a) \oplus_1 (f^{-1})(b))) = (f^{-1})(a) \oplus_1 (f^{-1})(b)$
3. From $f(1_1) = 1_2$ it follows that $(f^{-1})(1_2) = 1_1$ □

Theorem 3.35. *If $\langle A, \oplus_A, \odot_A \rangle$ and $\langle B, \oplus_B, \odot_B \rangle$ are fields with additive units 0_A, 0_B and multiplicative units 1_A,$1_B$and $f: A \rightarrow B$ a field homeomorphism then we have*

1. $f(0_A) = 0_B$
2. $\forall a \in A$ *we have* $f(-a) = -f(a)$
3. $\forall a \in A$ *with* $a \neq 0_A$ *we have* $f(a^{-1}) = (f(a))^{-1}$

Proof. As field homeomorphism $f: A \rightarrow B$ for the fields $\langle A, \oplus_A, \odot_A \rangle, \langle B, \oplus_B, \odot_B \rangle$ is automatically a group homeomorphism for the groups $\langle A, \oplus_A \rangle, \langle B, \oplus_B \rangle$ we have by 3.13 that (1) and (2) are valid. As for (3), if $a \in A$ with $a \neq 0_A$ then there exists a a^{-1} such that $a^{-1} \cdot a = 1_A$ hence $1_B = f(1_A) = f(a^{-1} \cdot a) = f(a^{-1}) \cdot f(a)$ proving that $f(a^{-1}) = (f(a))^{-1}$. □

Chapter 4
The Natural Numbers

4.1 Definition of the Natural Numbers

Recall the definition of successor (see 1.73), successor set (see 1.74) and the Axiom of Infinity (1.75).

Definition 4.1. *If A is a set then the successor of A is defined to be the set $s(A)=A\bigcup\{A\}$*

Definition 4.2. *A set A is a successor set if*

1. $\emptyset\in A$
2. *If* $X\in A\Rightarrow s(X)\in A$

Axiom 4.3. (Axiom of Infinity) *There exists a successor set*

We can now define the set of natural numbers $\mathbb{N}_0$ as follows

Definition 4.4. (Natural Numbers) *Let $\$=\{S: S\,\mathrm{is}\,a\,\mathrm{successor\,set}\}$ then $\mathbb{N}_0=\bigcap_{S\in\$}S$ is the set of natural numbers. In other words the set of Natural Numbers is the intersection of all successor sets. Note that there exists a successor set S and $\mathbb{N}_0\subseteq S$ so we have by 1.64 that $\mathbb{N}_0$ is a set.*

As for all successor sets S we have $\emptyset\in S, s(\emptyset)\in S, s(s(\emptyset))\in S,\ldots$ so thus $\emptyset, s(\emptyset)$, $s(s(\emptyset))\ldots\in\mathbb{N}_0$

Definition 4.5. *The successor function $s:\mathbb{N}_0\rightarrow\mathbb{N}_0$ is the function defined by $n\rightarrow s(n)$*

Definition 4.6. *We define the numbers*

1. $0=\emptyset\in\mathbb{N}_0$
2. $1=s(\emptyset)=\emptyset\bigcup\{\emptyset\}=\{\emptyset\}=\{0\}\in\mathbb{N}_0$
3. $2=s(1)=1\bigcup\{1\}=\{0\}\bigcup\{1\}=\{0,1\}=\{\emptyset,\{\emptyset\}\}\in\mathbb{N}_0$
4. $3=s(2)=2\bigcup\{2\}=\{0,1\}\bigcup\{2\}=\{0,1,2\}=\{\emptyset,\{\emptyset\},\{\emptyset,\{\emptyset\}\}\}$
5. ...

Definition 4.7. $\mathbb{N}=\mathbb{N}_0\backslash\{0\}$ *(the set of non-zero natural numbers)*

Theorem 4.8. *If $n\in\mathbb{N}_0\Rightarrow s(n)\neq 0$*

Proof. By definition $s(n) = n \bigcup \{n\} \Rightarrow n \in s(n) \Rightarrow s(n) \neq \emptyset = 0$ □

Theorem 4.9. *If* $n \in \mathbb{N}_0 \Rightarrow s(n) \in \mathbb{N}_0$

Proof. If $n \in \mathbb{N}_0$ then if S is a successor set we have that $n \in S \Rightarrow s(n) \in S \Rightarrow s(n) \in \bigcap_{S \in \mathbb{S}} S = \mathbb{N}_0$ □

Theorem 4.10. (Mathematical Induction) *If* $X \subseteq \mathbb{N}_0$ *has the following properties*

1. $0 \in X$
2. $n \in X \Rightarrow s(n) \in X$

then we have $X = \mathbb{N}_0$

Proof. By (1) and (2) we have that X is a successor set and thus by 4.4 we have $\mathbb{N}_0 \subseteq X$ so that from $X \subseteq \mathbb{N}_0$ we have $X = \mathbb{N}_0$ □

Theorem 4.11. *Let* $m, n \in \mathbb{N}_0$ *be natural numbers then if* $m \in s(n)$ *then* $m \in n \vee m = n$

Proof. As $s(n) = n \bigcup \{n\}$ then from $m \in s(n)$ we have either $m \in n \vee m \in \{n\}$ so we have either $m \in n \vee m = n$ □

Definition 4.12. *A set A is* ***transitive*** *if* $\forall x \in A \vdash x \subseteq A$

Theorem 4.13. $\forall n \in \mathbb{N}_0$ *we have that n is transitive*

Proof. We prove this by mathematical induction so let $S = \{n \in \mathbb{N}_0 | n \,\mathrm{is\,transitive}\} \subseteq \mathbb{N}_0$. We have then

1. $0 = \emptyset$ so $\forall x \in \emptyset \vdash x \subseteq \emptyset$ is satisfied vacuously and thus $0 \in S$
2. If $n \in S$ then $\forall m \in s(n)$ we have by the previous theorem (4.11)
   a. $m \in n \underset{n\,\mathrm{is\,transitive}}{\Rightarrow} m \subseteq n \subseteq n \bigcup \{n\} = s(n) \Rightarrow m \subseteq s(n)$
   b. $m = n \Rightarrow m \subseteq n \bigcup \{n\} = s(n) \Rightarrow m \subseteq s(n)$

Using mathematical induction we have then that $S = \mathbb{N}_0$ so if $n \in \mathbb{N}_0 \Rightarrow n \in S \Rightarrow n$ is transitive. □

Theorem 4.14. *If* $n \in \mathbb{N}_0 \Rightarrow n \neq s(n)$

Proof. We prove this by mathematical induction, so let $S = \{n \in \mathbb{N}_0 | n \neq s(n)\}$ then we have

1. $0 \neq \emptyset \bigcup \{\emptyset\} = s(0)$ [as $\emptyset \notin \emptyset = 0$]$\Rightarrow 0 \in S$
2. If $n \in S$ then $n \neq s(n)$ we prove now by contradiction that $s(n) \neq s(s(n))$. So assume that $s(n) = s(s(n))$ now as $s(s(n)) = s(n) \bigcup \{s(n)\}$ we have $s(n) \in s(s(n)) = s(n) \Rightarrow s(n) \in s(n) = n \bigcup \{n\} \underset{n \neq s(n)}{\Rightarrow} s(n) \in n \underset{4.13}{\Rightarrow} s(n) \subseteq n \Rightarrow n \bigcup \{n\} \subseteq n \subseteq n \bigcup \{n\} \Rightarrow n = n \bigcup \{n\} = s(n)$ contradicting $n \neq s(n)$. So we must have $s(n) \neq s(s(n)) \Rightarrow s(n) \in S$

Using 4.10 we have then that $S = \mathbb{N}_0 \Rightarrow$ if $n \in \mathbb{N}_0 \Rightarrow n \in S \Rightarrow n \neq s(n)$ □

Theorem 4.15. *If $n, m \in \mathbb{N}_0$ is such that $s(n) = s(m)$ then $n = m$*

Proof. If $s(n) = s(m)$ then as $n \in n \bigcup \{n\} = s(n) \Rightarrow n \in s(m)$ then using 4.11 we have either

1. $n \in m \underset{m \text{is transitive}}{\Rightarrow} n \subseteq m$. Now $m \in s(m) = s(n) \Rightarrow m \in s(n)$ and by 4.11 we have then
   a. $m \in n$ now we have by transitivity of n that $m \subseteq n$ and thus $n = m$
   b. $m = n$ in this case the theorem is proved.
2. $n = m$ in this case the theorem is proved. □

We have proved now that $\mathbb{N}_0$ full fills the **Peano** axiom's

Theorem 4.16. (Peano Axioms) $\mathbb{N}_0$ *satisfies the following so-called Peano Axioms*

1. $0 \in \mathbb{N}_0$ *(see 4.6)*
2. *If* $n \in \mathbb{N}_0 \Rightarrow s(n) \in \mathbb{N}_0$ *(see 4.9)*
3. $\forall n \in \mathbb{N}_0$ *we have* $s(n) \neq 0$ *(see 4.8)*
4. *If* $X \subseteq \mathbb{N}_0$ *is such that*
   a. $0 \in X$
   b. $n \in X \Rightarrow s(n) \in X$

   then $X = \mathbb{N}_0$ *(see 4.10)*
5. *If* $n, m \in \mathbb{N}_0$ *is such that* $s(n) = s(m)$ *then* $n = m$ *(see 4.15)*

Theorem 4.17. *If* $n \in \mathbb{N}_0 \wedge n \neq 0$ *then* $\exists! m \in \mathbb{N}_0$ *such that* $n = s(m)$

Proof. Define $A = \{n \in \mathbb{N}_0 | (n = 0) \vee (\exists! m \in \mathbb{N}_0 \vdash n = s(m))\}$ then we have

1. $0 \in A$
2. If $n \in A \Rightarrow n \in \mathbb{N}_0$ and $s(n) = s(n)$ further if $m \in \mathbb{N}_0$ such that $s(n) = s(m) \Rightarrow n = m \Rightarrow \exists! r \in \mathbb{N}_0\, [\text{just take } r = n] \vdash s(n) = s(r)$ and thus $s(n) \in a$

So if $n \in \mathbb{N}_0 \wedge n \neq 0 \Rightarrow n \in A \wedge n \neq 0 \Rightarrow \exists! m \in \mathbb{N}_0 \vdash n = s(m)$ □

4.2 Recursion

Theorem 4.18. (Recursion) *If A is a set, $a \in A$ and $f: A \to A$ is a function then there exists a unique function $\lambda: \mathbb{N}_0 \to A$ such that*

1. $\lambda(0) = a$
2. $\forall n \in \mathbb{N}_0 \vDash \lambda(s(n)) = f(\lambda(n))$

Proof. Let $\mathcal{G} = \{G | G \subseteq \mathbb{N}_0 \times A \text{such that} (0, a) \in G \text{and} \forall n \in \mathbb{N}_0 \vDash (n, x) \in G \Rightarrow (s(n), f(x)) \in G\}$. We have then that $\mathcal{G} \neq \emptyset$ as for $\mathbb{N}_0 \times A$ we have $(0, a) \in \mathbb{N}_0 \times A$ and if $(n, x) \in \mathbb{N}_0 \times A$ then as $s(n) \in \mathbb{N}_0 \wedge f(x) \in A \Rightarrow (s(n), f(x)) \in \mathbb{N}_0 \times A \Rightarrow \mathbb{N}_0 \times A \in \mathcal{G}$. We prove now that

$$\lambda = \bigcap_{G \in \mathcal{G}} G \in \mathcal{G}$$

Proof.

1. $\lambda \subseteq N \times A$ as $\forall G \in \mathcal{G}$ we have $G \subseteq \mathbb{N}_0 \times A$
2. $\forall G \in \mathcal{G}$ we have $(0, a) \in G \Rightarrow (0, a) \in \lambda$
3. If $(n, x) \in \lambda \Rightarrow \forall G \in \mathcal{G}$ we have $(n, x) \in G \Rightarrow (s(n), f(x)) \in G \Rightarrow (s(n), f(x)) \in \lambda$

So we have that $\lambda \in \mathcal{G}$ □

We shown now that $\text{dom}(\lambda) = \mathbb{N}_0$. We have

1. $\text{dom}(\lambda) \subseteq \mathbb{N}_0$
2. $0 \in \text{dom}(\lambda)$ as $(0, a) \in \lambda$
3. If $n \in \text{dom}(\lambda) \Rightarrow \exists x \in A \vdash (n, x) \in \lambda \Rightarrow (s(n), f(x)) \in \lambda \Rightarrow s(n) \in \text{dom}(\lambda)$

Using mathematical induction (see 4.10) we have then that $\text{dom}(\lambda) = \mathbb{N}_0$. We prove now using mathematical induction that λ is a partial function (from $\text{dom}(\lambda) = \mathbb{N}_0$ we have then that λ is a function). So let $S = \{n \in \mathbb{N}_0 | \exists! x \vdash (n, x) \in \lambda\} \subseteq \mathbb{N}_0$ then we have

1. $0 \in S$. We prove this by contradiction, so let $(0, a), (0, x) \in \lambda$ with $a \neq x \Rightarrow (0, a) \neq (0, x)$ Take then $\lambda' = \lambda \backslash (0, x) \subset \lambda$. We have then

   a. $(0, a) \in \lambda \underset{(0,a) \neq (0,x)}{\Rightarrow} (0, a) \in \lambda'$

   b. If $(n, y) \in \lambda' \underset{\lambda' \subseteq \lambda}{\Rightarrow} (n, y) \in \lambda \Rightarrow (s(n), f(y)) \in \lambda \underset{4.8 \Rightarrow s(n) \neq 0}{\Rightarrow} (s(n), f(y)) \in \lambda'$

   this proves that $\lambda' \in \mathcal{G}$ but then we have $\lambda \subseteq \lambda' \Rightarrow \lambda \subset \lambda'$ a contradiction. So we must have that $0 \in S$

2. If $n \in S \Rightarrow s(n) \in S$. Again we prove this by contradiction. So assume that $n \in S \Rightarrow \exists! x \vdash (n, x) \in \lambda \Rightarrow (s(n), f(x)) \in \lambda$, assume now that $\exists y \neq f(x) \vdash (s(n), y) \in \lambda$. Then we can form $\lambda' = \lambda \setminus (s(n), y)$ where because $y \neq f(x)$ we have $\lambda' \subset \lambda$. For λ' we have then that

   a. $(0, a) \in \lambda \underset{0 \neq s(n) \Rightarrow (0,a) \neq (s(n), y)}{\Rightarrow} (0, a) \in \lambda'$

   b. If $(m, z) \in \lambda' \Rightarrow (m, z) \in \lambda \Rightarrow (s(m), f(z)) \in \lambda$ we have then two sub cases

      i. $s(m) = s(n)$ then by 4.15 we have that $m = n \Rightarrow (m, z) = (n, z) \underset{n \in S \Rightarrow (n,z), (n,x) \in \lambda \Rightarrow z = x}{\Rightarrow} z = x \Rightarrow (s(m), f(z)) = (s(n), f(x)) \neq (s(n), y) \Rightarrow (s(m), f(z)) \in \lambda'$

ii. $s(m) \neq s(n)$ then $(s(m), f(z)) \neq (s(n), y) \Rightarrow (s(m), f(z)) \in \lambda'$

or in all cases $(s(m), f(z)) \in \lambda'$

This proves that $\lambda' \in \mathcal{G}$ but then we have $\lambda \subseteq \lambda' \underset{\lambda' \subset \lambda}{\Rightarrow} \lambda \subset \lambda$ a contradiction. So we must conclude that $\forall y \vdash (s(n), y) \in \lambda$ that $y = f(x)$ and thus that $\exists! y (= f(x)) \vdash (s(n), y) \in \lambda$ proving that $s(n) \in S$

Using mathematical induction we have then that $S = \mathbb{N}_0$ and thus that λ is a partial function and because we have already proved that $\mathrm{dom}(\lambda) = \mathbb{N}_0$ we have that λ: $\mathbb{N}_0 \rightarrow A$ is a function. By the fact that $\lambda \in \mathcal{G}$ we have also that

1. $(0, a) \in \lambda$
2. if $(n, x) \in \lambda \Rightarrow (s(n), f(x)) \in \lambda$

proving that λ is the function we search for. Now to prove that λ is unique suppose that there exists a λ' satisfying (1) and (2) above, define $T = \{n | \lambda(n) = \lambda'(n)\}$, then we have

1. From $\lambda(0) = a = \lambda'(0)$ that $0 \in T$
2. If $n \in T \Rightarrow \lambda(n) = \lambda'(n) = d \Rightarrow (n, d) \in \lambda \wedge (n, d) \in \lambda' \underset{(2)\, f \,\text{is a fucntion}}{\Rightarrow} (s(n), f(d)) \in \lambda \wedge (s(n), f(d)) \in \lambda' \Rightarrow \lambda(s(n)) = f(d) = \lambda'(s(n)) \Rightarrow s(n) \in T$

By using mathematical induction again we have then that $T = \mathbb{N}_0$ so if $n \in \mathbb{N}_0 \Rightarrow n \in T \Rightarrow \lambda(n) = \lambda'(n) \Rightarrow \lambda = \lambda'$ □

As a corollary of the above theorem we have the following

Corollary 4.19. *If A is a set, $a \in A$ and $f: A \rightarrow A$ is a injective function then there exists a unique function $\lambda: \mathbb{N}_0 \rightarrow A$ such that*

1. $\lambda(0) = a$
2. $\forall n \in \mathbb{N}_0 \vDash \lambda(s(n)) = f(\lambda(n))$
3. *If $a \notin f(A)$ then λ is injective*

Proof. By recursion (see 4.18) we have that there exists a unique function λ: $\mathbb{N}_0 \rightarrow A$ such that (1) and (2) are satisfied. Now if $a \notin f(A)$ then we must show that if $\lambda(m) = \lambda(n)$ then $m = n$. We do this by induction so let $n \in \mathbb{N}_0$ and define $S = \{m \in \mathbb{N}_0 | \text{if } n \in \mathbb{N}_0 \text{ is such that } \lambda(m) = \lambda(n) \Rightarrow m = n\}$ then we have

1. Then $0 \in S$ as we have the following if $\lambda(n) = \lambda(0)$
   a. $n = 0$ then $0 \in S$
   b. $n \neq 0$ then by 4.17 there exists a $k \in \mathbb{N}_0$ such that $n = s(k)$ but then $\lambda(n) = \lambda(0) = a$ implies $\lambda(s(k)) = a \Rightarrow f(\lambda(k)) = a$ which is false as $a \notin f(A)$ and thus $\lambda(n) = \lambda(0) \Rightarrow n = 0$ is true or $0 \in S$ again.
2. If $m \in S$ then if $k \in \mathbb{N}_0 \vdash \lambda(m) = \lambda(k) \Rightarrow m = k$. Suppose now that there exists a $n \in \mathbb{N}_0 \vdash \lambda(s(m)) = \lambda(n)$ then we have the following possibilities
   a. $n = 0$ then $a = \lambda(0) = \lambda(s(m)) = f(\lambda(m))$ is not allowed because $a \notin f(A)$ so $\lambda(s(m)) = \lambda(n)$ is false and thus $\lambda(s(m)) = \lambda(n) \Rightarrow s(m) = n$ is true and thus $s(m) \in S_n$

b. $n \neq 0$ then by 4.17 we have that $\exists k \in \mathbb{N}_0 \vdash n = s(k)$ so from $\lambda(s(m)) = \lambda(n)$ we have $\lambda(s(m)) = \lambda(s(k)) \Rightarrow f(\lambda(m)) = f(\lambda(k)) \underset{f\,\text{is}\,\text{injective}}{\Rightarrow} \lambda(m) = \lambda(k) \underset{m \in S}{\Rightarrow} m = k \Rightarrow s(m) = s(k) \Rightarrow s(m) = n \Rightarrow s(m) \in S$

Using 4.10 we have then that $S = \mathbb{N}_0$ and thus if there exists a $n \in \mathbb{N}_0$ and a $m \in \mathbb{N}_0$ with $s(m) = s(n) \Rightarrow n \in S \Rightarrow m = n$ proving injectivity. $\square$

Theorem 4.20. *Suppose A is a set $a \in A$ and $f: A \rightarrow A$ a function, then there exists a* ***unique*** *function $\varphi: \mathbb{N}_0 \rightarrow A$ such that $\varphi(0) = a$ and $\varphi \circ s = f \circ \varphi$.*

Proof. Define the function $\sigma: \mathbb{N}_0 \times A \rightarrow A$ by $(n, x) \rightarrow (s(n), f(x))$. Given $a \in A$ we define a subset $R \subseteq \mathbb{N}_0 \times A$ to be a-closed if the following is true

1. $(0, a) \in R$
2. $\sigma(R) \subseteq R$

We have that $\mathbb{N}_0 \times A$ is a-closed for

1. $(0, a) \in \mathbb{N}_0 \times A$
2. If $(m, y) \in \sigma(\mathbb{N}_0 \times A) \Rightarrow \exists (n, x) \in \mathbb{N}_0 \times A \vdash (m, y) = (s(n), f(x)) \in \mathbb{N}_0 \times A \Rightarrow \sigma(\mathbb{N}_0 \times A) \subseteq \mathbb{N}_0 \times A$.

If $\mathcal{R}$ is a set of a-closed subsets of $\mathbb{N}_0 \times A$ then we have

1. $\forall R \in \mathcal{R}$ we have $(0, a) \in R \Rightarrow (0, a) \in \bigcap_{R \in \mathcal{R}} R$
2. if $y \in \sigma(\bigcap_{R \in \mathcal{R}} R) \Rightarrow \exists x \in \bigcap_{R \in \mathcal{R}} R \vdash y = \sigma(x) \Rightarrow \exists x \vdash (\forall R \in \mathcal{R} \vdash x \in R) \wedge y = \sigma(x) \Rightarrow \forall R \in \mathcal{R} \vdash y \in \sigma(R) \Rightarrow y \in \bigcap_{R \in \mathcal{R}} \sigma(R) \Rightarrow \sigma(\bigcap_{R \in \mathcal{R}} R) \subseteq \bigcap_{R \in \mathcal{R}} \sigma(R) \underset{\sigma(R) \subseteq R}{\Rightarrow} \sigma(\bigcap_{R \in \mathcal{R}} R) \subseteq \bigcap_{R \in \mathcal{R}} R$

proving that $\bigcap_{R \in \mathcal{R}} R$ is also a-closed. Define now the set $\mathcal{R}_a = \{R \in \mathcal{P}(\mathbb{N}_0 \times A) | R\,\text{is a-closed}\}$ set (because $\subseteq \mathcal{P}(\mathbb{N}_0 \times A)$ which is set because $\mathbb{N}_0 \times A$ is a set and 1.69) and $R_a = \bigcap_{R \in \mathcal{R}_a} R$ then we have by the above that R_a is a-closed and is further the smallest a-closed set. We have then

a) $(0, a) \in \{(0, a)\} \bigcup \sigma(R_a)$
b) $\sigma(\{(0, a)\} \bigcup \sigma(R_a)) \subseteq \sigma(\{(0, a)\}) \bigcup \sigma(\sigma(R_a)) \subseteq_{\sigma(0,a) \in R_a} R_a \bigcup \sigma(\sigma(R_a)) \subseteq_{\sigma(R_a) \subseteq R_a} R_a \bigcup \sigma(R_a) \subseteq_{\sigma(R_a) \subseteq R_a} R_a \bigcup R_a = R_a$

proving that $\{(0, a)\} \bigcup R_a$ is a-closed and then from the minimality of R_a we have

$$R_\alpha \subseteq \{(0, a)\} \bigcup \sigma(R_a)$$

As $\{(0, a)\} \subseteq R_a$ and $\sigma(R_a) \subseteq R_a$ we have

$$\{(0, a)\} \bigcup \sigma(R_a) \subseteq R_a$$

and thus we finally have

$$R_a = \{(0, a)\} \bigcup \sigma(R_a)$$

We say now that $n \in \mathbb{N}_0$ is a-paired-up iff

a) There exists $y \in A$ such that $(n, y) \in R_a$

b) If also $(n, y') \in R_a \Rightarrow y = y'$

Let's define now the set of a-paired-up numbers $P_a = \{n \in \mathbb{N}_0 | n \text{ is a-paired-up}\} \subseteq \mathbb{N}_0$ then we proof that

1. $0 \in P_a$

   Proof. First $(0, a) \in R_a$. Second if $(0, a') \in R_a$ with $a \neq a' \Rightarrow (0, a) \neq (0, a') \underset{(0,a') \in R_a = \{(0,a)\} \bigcup \sigma(R_a)}{\Rightarrow} (0, a') \in \sigma(R_a)$ so there exists a $(n, a'') \in R_a$ with $(0, a') = \sigma(n, a'') = (s(n), f(a''))$ giving $0 = s(n)$ a contradiction. So we must conclude that $a = a'$. □

2. If $n \in P_a \Rightarrow s(n) \in P_a$

   Proof. As $n \in P_a$ then n is a-paired-up so there exists a **unique** $y \in A$ such that $(n, y) \in R_a \underset{R_a \text{ is a-closed}}{\Rightarrow} \sigma(n, y) \in R_a \Rightarrow (s(n), f(y)) \in R_a$. Suppose now that $(s(n), y') \in R_a = \{(0, a)\} \bigcup \sigma(R_a) \underset{s(n) \neq 0}{\Rightarrow} (s(n), y') \in \sigma(R_a) \Rightarrow \exists (m, c) \in R_a \vdash (s(n), y') = \sigma(m, c) = (s(m), f(c)) \Rightarrow s(n) = s(m) \wedge y' = f(c) \underset{4.15}{\Rightarrow} n = m \wedge y' = f(c) \Rightarrow (n, c) \in R_a \wedge y' = f(c) \underset{(n, y) \in R_a}{\Rightarrow} (n, c), (n, y) \in R_a \wedge y' = f(c) \underset{n \text{ is a-paired-up}}{\Rightarrow} c = y \wedge y' = f(c) \Rightarrow y' = f(y)$. So $s(n)$ is a-paired-up and thus $s(n) \in P_a$ □

Using mathematical induction (see 4.10) we have then that $P_a = \mathbb{N}_0$. Define now $\varphi: \mathbb{N}_0 \rightarrow A$ with $\varphi = R_a$ then we have that if $n \in \mathbb{N}_0 \Rightarrow n \in P_a$ so there exists a $y \in A \vdash (n, a) \in R_a = \varphi$ and thus $\mathrm{dom}(\varphi) = \mathbb{N}_0$. Second if $(n, y), (n, y') \in \varphi$ then as $n \in \mathbb{N}_0 \Rightarrow n \in P_a$ we have that n is a-paired-up and thus $y = y'$. So we conclude that $\varphi: \mathbb{N}_0 \rightarrow A$ is indeed a function. Now as R_a is a-closed we have $(0, a) \in R_a = \varphi \Rightarrow \varphi(0) = a$ and thus

$$\varphi(0) = a$$

Second if $n \in \mathbb{N}_0$ then $n \in P_a$ so there exists a $y \in A$ such that $(n, y) \in R_a = \varphi \Rightarrow y = \varphi(n)$ then as R_a is a-closed we have $(s(n), f(y)) \in R_a = \varphi$ and thus $(\varphi \circ s)(n) = \varphi(s(n)) = f(y) = f(\varphi(n)) = (f \circ \varphi)(n)$ proving that

$$\varphi \circ s = f \circ \varphi$$

So we have found that the function $\varphi: \mathbb{N}_0 \rightarrow A$ satisfies the requirements of the theorem. Let's now prove that it is unique. Assume that there exists also a φ': $\mathbb{N}_0 \rightarrow A$ satisfying

$$\begin{aligned} \varphi'(0) &= a \\ \varphi' \circ s &= f \circ \varphi \end{aligned}$$

Take then $E = \{n \in \mathbb{N}_0 | \varphi(n) = \varphi'(n)\}$ then we have

1. $\varphi(0) = a = \varphi'(0) \Rightarrow 0 \in A$

2. If $n \in E$ then for $s(n)$ we have $\varphi(s(n)) = (\varphi \circ s)(n) = (f \circ \varphi)(n) = f(\varphi(n)) \underset{n \in E \Rightarrow \varphi(n) = \varphi'(n)}{=} f(\varphi'(n)) = (f \circ \varphi')(n) = (\varphi' \circ s)(n) = \varphi'(s(n))$ and thus $s(n) \in E$

Using mathematical induction (see 4.10) we have then that $E = \mathbb{N}_0$ so $\forall n \in \mathbb{N}_0$ we have $n \in E \Rightarrow \varphi(n) = \varphi'(n)$ and thus $\varphi = \varphi'$

□

This theorem allows us to iterate the application of a function in the following way: If $a \in A$ and $f: A \rightarrow A$ is a function then there exists a function $\varphi: \mathbb{N}_0 \rightarrow A$ such that

$$\begin{array}{rcl} \varphi(0) & = & a \\ \varphi(1) = \varphi(s(0)) & = & f(\varphi(0)) = f(a) \\ \varphi(2) = \varphi(s(1)) & = & f(\varphi(1)) = f(f(a)) \\ \varphi(3) = \varphi(s(2)) & = & f(\varphi(2)) = f(f(f(a))) \\ & \ldots & \\ \varphi(n) & = & \overbrace{f(f(....(f(a))))}^{n \text{times}} \end{array}$$

As application of the above theorem we have the following iteration theorem

Theorem 4.21. (Iteration) *Let A be a set and $f: A \rightarrow A$ be a function. Then $\forall n \in \mathbb{N}_0$ there exists a function $(f)^n: A \rightarrow A$ such that*

1. $(f)^0 = 1_A$
2. $(f)^{s(n)} = f \circ (f)^n$

Proof. If $A = \emptyset$ then $\emptyset: \emptyset \rightarrow \emptyset$ then as $1_\emptyset: \emptyset \rightarrow \emptyset$ is also the empty function $\emptyset$ then $(\emptyset)^n = \emptyset$ satisfies (1), (2). If $A \neq \emptyset$ then let $n \in \mathbb{N}_0$ and $a \in A$ and use the previous theorem 4.20 to find a function $\varphi_a: \mathbb{N}_0 \rightarrow A$ such that $\varphi_a(0) = a$ and $\varphi_a \circ s = f \circ \varphi_a$. Define now $(f)^n: A \rightarrow A$ by $a \rightarrow \varphi_a(n)$ then we have

1. $\forall a \in A$ that $(f)^0(a) = \varphi_a(0) = a \Rightarrow (f)^0 = i_A$
2. $\forall a \in A$ then $(f)^{s(n)}(a) = \varphi_a(s(n)) = f(\varphi_a(n)) = f((f)^n(a)) = (f \circ (f)^n)(a) \Rightarrow (f)^{s(n)} = f \circ (f)^n$

□

As a illustration of this iteration let $f: A \rightarrow A$ then we have

$$\begin{array}{rcl} (f)^0 & = & 1_A \\ (f)^1 = (f)^{s(0)} & = & f \circ (f)^0 = f \circ i_A = f \\ (f)^2 = (f)^{s(1)} & = & f \circ (f)^1 = f \circ f \\ (f)^3 = (f)^{s(2)} & = & f \circ (f)^2 = f \circ f \circ f \\ & \ldots & \\ (f)^n & = & \overbrace{f \circ \ldots \circ f}^{n \text{times}} \end{array}$$

Example 4.22. Let $\langle A, \oplus\rangle$ be a group and $a \in A$ define then $\oplus_a: A \rightarrow A$ by $x \rightarrow \oplus_a(x) = x \oplus a$ we define then given $n \in \mathbb{N}_0$ $a\langle\oplus\rangle n = (\oplus_a)^n(\nu)$ where ν is the neutral element in the group . We have then that

$$\begin{aligned} a\langle\oplus\rangle 0 &= (\oplus_a)^0(\nu) = i_A(\nu) = \nu \\ a\langle\oplus\rangle s(n) &= ((\oplus_a)^n(\nu)) \oplus a \\ &= (a\langle\oplus\rangle n) \oplus a \end{aligned}$$

Sometimes we consider a group to be additive or multiplicative this is either noted as $\langle A, +\rangle$ with neutral element 0 or $\langle A, \cdot\rangle$ with neutral element 1 then we note $a\langle+\rangle n$ as $a \cdot n$ as and $a\langle\cdot\rangle n$ as a^n we have then

1. Additive group $\langle A, +\rangle$ with neutral element 0

$$\begin{aligned} a \cdot 0 &= 0 \\ a \cdot s(n) &= (a \cdot n) + a \end{aligned}$$

so we have

$$\begin{aligned} a \cdot 0 &= 0 \\ a \cdot 1 &= a \cdot s(0) = (a \cdot 0) + a = 0 + a = a \\ a \cdot 2 &= a \cdot s(1) = (a \cdot 1) + a = a + a \\ a \cdot 3 &= a \cdot s(2) = (a \cdot 2) + a = (a + a) + a \\ &\ldots \end{aligned}$$

2. Multiplicative group $\langle A, \cdot\rangle$ with neutral element 1

$$\begin{aligned} a^0 &= 1 \\ a^{s(n)} &= (a^n) \cdot a \end{aligned}$$

so we have

$$\begin{aligned} a^0 &= 1 \\ a^1 &= a^{s(0)} = (a^0) \cdot a = 1 \cdot a = a \\ a^2 &= a^{s(1)} = (a^1) \cdot a = a \cdot a \\ a^3 &= a^{s(2)} = (a \cdot a) \cdot a \\ &\ldots \end{aligned}$$

Definition 4.23. *If $\langle F, +, \cdot\rangle$ is a field with multiplicative unit u and additive neutral element e, let $n \in \mathbb{N}$ define take then $f: F \rightarrow F$ defined by $x \rightarrow x + u$ then $n \cdot u = (f)^n(e)$.*

Example 4.24. If $\langle F, +, \cdot\rangle$ is a field with multiplicative unit u then we have

$$\begin{aligned} 0 \cdot u = (f)^0(e) &= 1_F(e) = e \\ 1 \cdot u = (f)^1(e) &= f^{(s(0))}(e) = f(f^0(e)) = f(e) = e + u = u \\ 2 \cdot u = (f)^2(e) &= f^{s(1)}(e) = f(f^1(e)) = f(u) = u + u \\ 3.u = (f)^3(e) &= f^{s(2)}(e) = f(f^2(e)) = f(u + u) = u + u + u \\ &\ldots \end{aligned}$$

Definition 4.25. *If $\langle F, +, \cdot \rangle$ is a field with multiplicative unit u and neutral element e then F is of characterization zero if $\forall n \in \mathbb{N}$ we have $n \cdot u \neq e$*

Theorem 4.26. (Recursion on $\mathbb{N}_0$ - Step Form) *Let A be a set, $a \in A$ and g: $\mathbb{N}_0 \times A \rightarrow A$ a function. Then there exists a unique function $f: \mathbb{N}_0 \rightarrow A$ satisfying*

$$\begin{aligned} f(0) &= a \\ \forall n \in \mathbb{N}_0 \text{ we have } f(s(n)) &= g(n, f(n)) \end{aligned}$$

Proof. First define the projection functions π_1 and π_2 as follows

$$\begin{aligned} &\pi_1: \mathbb{N}_0 \times A \rightarrow \mathbb{N}_0 \ \text{ by } \ (n, x) \rightarrow \pi_1(n, x) = n \\ &\pi_2: \mathbb{N}_0 \times A \rightarrow \mathbb{N}_0 \ \text{ by } \ (n, x) \rightarrow \pi_2(n, x) = x \end{aligned}$$

we haven then $\forall (n, x) \in \mathbb{N}_0 \times A$ that $(n, x) = (\pi_1(n, x), \pi_2(n, x))$. Next define γ by

$$\gamma: \mathbb{N}_0 \times A \rightarrow \mathbb{N}_0 \times A \ \text{ by } \ (n, x) \rightarrow \gamma(n, x) = (s(n), g(n, x))$$

By using 4.21 we have $\forall n \in \mathbb{N}_0$ the existence of a function $(\gamma)^n: \mathbb{N}_0 \times A \rightarrow \mathbb{N}_0 \times A$ such that

$$\begin{aligned} (\gamma)^0 &= 1_{\mathbb{N}_0 \times A} \\ (\gamma)^{s(n)} &= \gamma \circ (\gamma)^n \end{aligned}$$

We prove now by induction that $\forall n \in \mathbb{N}_0$ we have $\pi_1((\gamma)^n(0, a)) = n$. So define $B = \{n \in \mathbb{N}_0 | \pi_1((\gamma)^n(0, a)) = n\}$ then we have

1. $\pi_1((\gamma)^0(0, a)) = \pi_1(i_{\mathrm{N} \times A}(0, a)) = \pi_1(0, a) = 0$ and thus we have that $0 \in B$.
2. If $n \in B$ then $\pi_1((\gamma)^{s(n)}(0, a)) = \pi_1(\gamma((\gamma)^n(0, a))) = \pi_1(\gamma(\pi_1((\gamma)^n(0, a)), \pi_2((\gamma)^n(0, a)))) \underset{n \in B}{=} \pi_1(\gamma(n, \pi_2((\gamma)^n(0, a)))) = \pi_1(s(n), g(n, \pi_2((\gamma)^n(0, a)))) = s(n)$ and thus $s(n) \in B$

Using induction (see 4.10) we have then that $B = \mathbb{N}_0$ and thus proved our statement.

Define now $f: \mathbb{N}_0 \rightarrow A$ by $n \rightarrow f(n) = \pi_2((\gamma)^n(0, a))$ then we have

1. $f(0) = \pi_2((\gamma)^0(0, a)) = \pi_2(i_{N \times A}(0, a)) = \pi_2(0, a) = a$
2. If $n \in \mathbb{N}_0$ then $f(s(n)) = \pi_2((\gamma)^{s(n)}(0, a)) = \pi_2(\gamma((\gamma)^n(0, a)) = \pi_2(\gamma(\pi_1((\gamma)^n(0, a)), \pi_2((\gamma)^n(0, a)))) = \pi_2(\gamma(n, \pi_2((\gamma)^n(0, a)))) = \pi_2(s(n), g(n, \pi_2((\gamma)^n(0, a)))) = g(n, \pi_2((\gamma)^n(0, a))) = g(n, f(n))$

proving that f is the function we search for . We are thus left with proving uniqueness. So assume that there exists a $f': \mathbb{N}_0 \rightarrow A$ such that

$$\begin{aligned} f'(0) &= a \\ f'(s(n)) &= g(n, f'(n)) \end{aligned}$$

and define then $C = \{n \in \mathbb{N}_0 | f(n) = f'(n)\}$ then we have

1. $f(0) = a = f'(0)$ and thus $0 \in C$
2. If $n \in \mathbb{N}_0$ then $f(s(n)) = g(n, f(n)) = g(n, f'(n)) = f'(s(n))$ and thus $s(n) \in C$

Using mathematical induction (see 4.10) we have that $C = \mathbb{N}_0$ and thus $f = f'$ $\square$

4.3 Arithmetic of the natural numbers

Definition 4.27. (Addition) *Let $m, n \in \mathbb{N}_0$ then we define $+: \mathbb{N}_0 \times \mathbb{N}_0 \rightarrow \mathbb{N}_0$ by $(n, m) \rightarrow n+m = (s)^m(n)$ (here s is the successor function)*

Example 4.28. If $n \in \mathbb{N}_0$ then we have

$$\begin{array}{rcl} n+0 & = & (s)^0(n) = i_{\mathbb{N}_0}(n) = n \\ n+1 & = & (s)^1(n) = (s)^{s(0)}(n) = (s \circ (s)^0)(n) = s(i_{\mathbb{N}_0}(n)) = s(n) \\ n+2 & = & (s)^2(n) = (s)^{s(1)}(n) = (s \circ (s)^1)(n) = s(s(n)) \\ & & \ldots \end{array}$$

Using the definition of + we immediately have

Theorem 4.29. *If $n \in \mathbb{N}_0$ then $n+0 = 0+n = n$*

Proof.

1. $n+0 = (s)^0(n) = 1_{\mathbb{N}_0}(n) = n$
2. We prove this by mathematical induction, define $S = \{n \in \mathbb{N}_0 | 0+n = n\} \subseteq \mathbb{N}_0$ then we have
   a. $0+0 \underset{(1)}{=} 0 \Rightarrow 0 \in S$
   b. If $n \in S \Rightarrow 0+n = n$ now for $s(n)$ we have $0+s(n) = (s)^{s(n)}(0) = (s \circ s^{(n)})(0) = s(s^{(n)}(0)) = s(0+n) = s(n) \Rightarrow s(n) \in S$

Using mathematical induction 4.10 we have $S = \mathbb{N}_0$ and thus if $n \in \mathbb{N}_0 \Rightarrow n \in S \Rightarrow 0+n = n$ □

Theorem 4.30. *If $n \in \mathbb{N}_0$ then $s(n) = n+1 = 1+n$*

Proof.

1. $n+1 = (s)^1(n) = (s)^{s(0)}(n) = (s \circ (s)^0)(n) = (s \circ i_{\mathbb{N}_0})(n) = s(n)$
2. $1+n = s(n)$ is proved by induction, so define $S = \{n \in \mathbb{N}_0 | 1+n = s(n)\}$ then we have
   a. $1+0 \underset{4.29}{=} 1 = s(0) \Rightarrow 0 \in S$
   b. If $n \in S \Rightarrow 1+n = s(n)$. Now $1+s(n) = (s)^{s(n)}(1) = (s \circ (s)^n)(1) = s((s)^n(1)) = s(1+n) = s(s(n))$ and thus $s(n) \in S$

from 4.10 we have then that $S = \mathbb{N}_0 \Rightarrow n \in \mathbb{N}_0 \Rightarrow n \in S \Rightarrow 1+n = s(n)$ □

Lemma 4.31. *If $n, m \in \mathbb{N}_0$ then $n+s(m) = s(n+m)$*

Proof. $n+s(m) = (s)^{s(m)}(n) = (s \circ (s)^m)(n) = s((s)^m(n)) = s(n+m)$ □

Theorem 4.32. (Associativity) *If $n, m, k \in \mathbb{N}_0$ then $(n+m)+k = n+(m+k)$*

Proof. The proof is by mathematical induction given $n, m \in \mathbb{N}_0$ define $S_{n,m} = \{k \in \mathbb{N}_0 | (n+m)+k = n+(m+k)\}$ then we have

1. $(n+m)+0 \underset{4.29}{=} n+m \underset{4.29}{=} n+(m+0) \Rightarrow 0 \in S_{n,m}$
2. If $k \in S_{n,m} \Rightarrow (n+m)+k = n+(m+k)$. We have then $(n+m)+s(k) \underset{4.31}{=} s((n+m)+k) = s(n+(m+k)) \underset{4.31}{=} n+s(m+k) \underset{4.31}{=} n+(m+s(k))$ and thus $s(k) \in S_{n,m}$

So if $n, m, k \in \mathbb{N}_0$ then $k \in S_{n,m} \Rightarrow (n+m)+k = n+(m+k)$ □

Theorem 4.33. (Commutativity) *If $n, m \in \mathbb{N}_0$ then $n+m = m+n$*

Proof. Again we prove this by induction. So let $n \in \mathbb{N}_0$ define then $S_n = \{m \in \mathbb{N}_0 | n+m = m+n\}$ then we have

1. $n+0 = n = 0+n$ (see 4.29) and thus $0 \in S_n$
2. If $m \in S_n \Rightarrow n+m = m+n$. We have then $n+s(m) \underset{4.31}{=} s(n+m) = s(m+n) \underset{4.30}{=} 1+(m+n) \underset{4.32}{=} (1+n)+m \underset{4.30}{=} s(n)+m$ so we have that $s(m) \in S_n$

Using 4.10 we have then that $S_n = \mathbb{N}_0$. So if $n, m \in \mathbb{N}_0 \Rightarrow m \in S_n \Rightarrow n+m = m+n$ □

Corollary 4.34. *$\langle \mathbb{N}_0, + \rangle$ forms a abelian semi-group*

Proof. This follows from 4.32,4.29 and 4.33 □

Definition 4.35. (Multiplication) *Given $n \in \mathbb{N}_0$ define $\alpha_n: \mathbb{N}_0 \to \mathbb{N}_0$ defined by $m \to \alpha_n(m) = n+m$. Then we define $\cdot: \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$ by $(n, m) \to n \cdot m = (\alpha_n)^m(0)$*

Example 4.36. We have the following examples to see how multiplication works by repeating summation

$$\begin{aligned}
2 \cdot 0 &= (\alpha_2)^0(0) = i_{\mathbb{N}_0}(0) = 0 \\
2 \cdot 1 &= (\alpha_2)^1(0) = (\alpha_2)^{s(0)}(0) = (\alpha_2 \circ (\alpha_2)^0)(0) = \alpha_2(0) = 2+0 = 2 \\
2 \cdot 2 &= (\alpha_2)^2(0) = (\alpha_2)^{s(1)}(0) = (\alpha_2((\alpha_2)^1(0))) = \alpha_2(2) = 2+2 = 4
\end{aligned}$$

Theorem 4.37. (Absorbing Element) *If $n \in \mathbb{N}_0$ then $n \cdot 0 = 0 = 0 \cdot n$*

Proof.

1. $n \cdot 0 = (\alpha_n)^0(0) = 1_{\mathbb{N}_0}(0) = 0$
2. We prove by induction that $0 \cdot n = 0$ so define $S = \{n \in \mathbb{N}_0 | 0 \cdot n = 0\} \subseteq \mathbb{N}_0$ then we have
   a. $0 \cdot 0 \underset{(1)}{=} 0 \Rightarrow 0 \in S$
   b. If $n \in S$ then $0 \cdot n = 0$ Now $0 \cdot s(n) = (\alpha_0)^{s(n)}(0) = (\alpha_0 \circ (\alpha_0)^n)(0) = \alpha_0((\alpha_0)^n(0)) = \alpha_0(0 \cdot n) = \alpha_0(0) = 0+0 \underset{4.29}{=} 0$ and thus $s(n) \in S$

   proving by 4.10 that $S = \mathbb{N}_0$ □

Theorem 4.38. (Neutral Element) *If* $n \in \mathbb{N}_0$ *then* $n \cdot 1 = n = 1 \cdot n$

Proof.

1. $n \cdot 1 = (\alpha_n)^1(0) = (\alpha_n)^{s(0)}(0) = \alpha_n((\alpha_n)^0(0)) = \alpha_n(0) = n+0 \underset{4.29}{=} n$
2. We prove by induction that $1 \cdot n = n$ so define $S = \{n \in \mathbb{N}_0 | 1 \cdot n = n\}$ then we have
   a. $1 \cdot 0 \underset{4.37}{=} 0 \Rightarrow 0 \in S$
   b. If $n \in S \Rightarrow 1 \cdot n = n$ Now $1 \cdot s(n) = (\alpha_1)^{s(n)}(0) = \alpha_1((\alpha_1)^n(0)) = \alpha_1(1 \cdot n) = \alpha_1(n) = 1+n \underset{4.30}{=} s(n)$ so $s(n) \in S$

   Using 4.10 we have then $S = \mathbb{N}_0$ □

Theorem 4.39. *If* $n, m \in \mathbb{N}_0$ *then* $n \cdot s(m) = n+n \cdot m \underset{4.33}{=} n \cdot m+n$

Proof. $n \cdot s(m) = (\alpha_n)^{s(m)}(0) = \alpha_n((\alpha_n)^m(0)) = \alpha_n(n \cdot m) = n+n \cdot m$ □

Theorem 4.40. (Distributivity) $\forall n, m, k \in \mathbb{N}_0$ *we have* $(n+m) \cdot k = n \cdot k + m \cdot k$

Proof. If $n, m \in \mathbb{N}_0$ define $S_{n,m} = \{k \in \mathbb{N}_0 | (n+m) \cdot k = n \cdot k+m \cdot k\}$ then we have

1. $(n+m) \cdot 0 \underset{4.37}{=} 0 \underset{4.29}{=} 0+0 \underset{4.37}{=} n \cdot 0 + m \cdot 0 \Rightarrow 0 \in S_{n,m}$
2. Assume that $k \in S_{n,m}$ then $(n+m) \cdot k = n \cdot k+m \cdot k$. Now $(n+m) \cdot s(k) \underset{4.39}{=} (n+m) \cdot k+(n+m) = (n \cdot k + m \cdot k)+(n+m) \underset{4.32}{=} (n \cdot k + n) + (m \cdot k + m) \underset{4.39}{=} n \cdot s(k) + m \cdot s(k) \Rightarrow s(k) \in S_{n,m}$

If now $n, m, k \in \mathbb{N}_0 \Rightarrow k \in S_{n,m} \Rightarrow (n+m) \cdot k = n \cdot k + m \cdot k$ □

Theorem 4.41. (Commutativity) *If* $n, m \in \mathbb{N}_0$ *then* $n \cdot m = m \cdot n$

Proof. We prove this by induction, let $S_n = \{m \in \mathbb{N}_0 | n \cdot m = m \cdot n\}$ then we have

1. Using 4.37 we have $n \cdot 0 = 0 = 0 \cdot n \Rightarrow 0 \in S_n$
2. If $m \in S_n \Rightarrow n \cdot m = m \cdot n$. Now $n \cdot s(m) = n + n \cdot m \underset{m \in S_n}{=} n+m \cdot n \underset{4.38}{=} 1 \cdot n+m \cdot n \underset{4.40}{=} (1+m) \cdot n \underset{4.30}{=} s(m) \cdot n \Rightarrow s(m) \in S_n$

This proves by 4.10 that $S_n = \mathbb{N}_0$ and thus if $n, m \in \mathbb{N}_0 \Rightarrow m \in S_n \Rightarrow n \cdot m = m \cdot n$ □

Theorem 4.42. (Associativity) *If* $n, m, k \in \mathbb{N}_0$ *then* $(n \cdot m) \cdot k = n \cdot (m \cdot k)$

Proof. Given $n, m \in \mathbb{N}_0$ define $S_{n,m} = \{k \in \mathbb{N}_0 | (n \cdot m) \cdot k = n \cdot (m \cdot k)\}$ then we have

1. $(n \cdot m) \cdot 0 \underset{4.37}{=} 0 \underset{4.37}{=} n \cdot 0 \underset{4.37}{=} n \cdot (m \cdot 0) \Rightarrow 0 \in S_{n,m}$
2. If $k \in S_{n,m}$ then $(n \cdot m) \cdot k = n \cdot (m \cdot k)$. Now $(n \cdot m) \cdot s(k) \underset{4.39}{=} (n \cdot m) \cdot k + n \cdot m = n \cdot (m \cdot k) + n \cdot m \underset{4.41}{=} (m \cdot k) \cdot n + m \cdot n \underset{4.40}{=} (m \cdot k + m) \cdot n \underset{4.39}{=} (m \cdot s(k)) \cdot n \underset{4.41}{=} n \cdot (m \cdot s(k)) \Rightarrow s(k) \in S_{n,m}$

Using mathematical induction 4.10 we have $S_{n,m} = \mathbb{N}_0$ so if $n, m, k \in \mathbb{N}_0 \Rightarrow k \in S_{n,m} \Rightarrow (n \cdot m) \cdot k = n \cdot (m \cdot k)$ □

Corollary 4.43. *$\langle \mathbb{N}_0, .\rangle$ is a abelian semi-group*

Proof. This follows from 4.41,4.42 and 4.38. □

Theorem 4.44. *If $n, m, k \in \mathbb{N}_0$ is such that $n+k = m+k \Rightarrow n = m$, (and thus we have $n = m \Leftrightarrow n+k = m+k$)*

Proof. We prove this by induction so let $S_{n.m} = \{k \in \mathbb{N}_0 | n+k = m+k \Rightarrow n = m\}$ then we have

1. $n+0 = m+0 \underset{4.29}{\Rightarrow} n = m \Rightarrow 0 \in S_{n,m}$
2. If $k \in S_{n,m} \Rightarrow n+k = m+k \Rightarrow n = m$. For $s(k)$ if we have $n+s(k) = m+s(k) \underset{4.31}{\Rightarrow} s(n+k) = s(m+k) \underset{4.15}{\Rightarrow} n+k = m+k \Rightarrow n = m \Rightarrow s(k) \in S_{n,m}$

So if $n, m, k \in \mathbb{N}_0 \Rightarrow k \in S_{n,m} \Rightarrow n+k = m+k \Rightarrow n = m$ □

4.4 Order relation on the natural numbers

Definition 4.45. *Define the relation $\leqslant \in \mathbb{N}_0 \times \mathbb{N}_0$ by $\leqslant = \{(n,m) \in \mathbb{N}_0 \times \mathbb{N}_0 | n \in m \vee n = m\}$ so we say that $n \leqslant m$ iff $n \in m \vee n = m$. If $n \leqslant m$ and $n \neq m$ then we say that $n < m$.*

Theorem 4.46. *$\langle \mathbb{N}_0, \leqslant\rangle$ is a partially ordered set*

Proof.

1. **(reflectivity)** If $m \in \mathbb{N}_0 \Rightarrow m = m \Rightarrow m \leqslant m$
2. **(anti-symmetry)** If $n, m \in \mathbb{N}_0$ and $n \leqslant m \wedge m \leqslant n \Rightarrow (n \in m \vee n = m) \wedge (m \in n \vee n = m) \Rightarrow (n \in m \wedge m \in n) \vee (n = m)$ so we have the following cases
   a. $n = m \Rightarrow$ anti-symmetry is proved
   b. $n \in m \wedge m \in n$ then by 4.13 $n \subseteq m \wedge m \subseteq n \Rightarrow n = m$ proving again anti-symmetry
3. **(transitivity)** If $n \leqslant m \wedge m \leqslant k$ then we have the following possibilities
   a. $n \in m \wedge m \in k$ then by 4.13 $n \subseteq m \wedge m \subseteq k \Rightarrow n \subseteq k \Rightarrow n \leqslant k$
   b. $n \in m \wedge m = k$ then by 4.13 $n \subseteq m \wedge m = k \Rightarrow n \subseteq k \Rightarrow n \leqslant k$
   c. $n = m \wedge m \in k$ then by 4.13 $n = m \wedge m \subseteq k \Rightarrow n \subseteq k \Rightarrow n \leqslant k$
   d. $n = m \wedge m = k$ then $n = k \Rightarrow n \leqslant k$

   So in all cases we have $n \leqslant k$ proving transitivity. □

Theorem 4.47. *If $n \in \mathbb{N}_0 \Rightarrow 0 \leqslant n$*

Proof. We prove this by mathematical induction, so let $S = \{n \in \mathbb{N}_0 | 0 \leqslant n\} \subseteq \mathbb{N}_0$ then we have

1. $0 = 0 \Rightarrow 0 \leqslant 0 \Rightarrow 0 \in S$

2. Suppose $n \in S$ then $0 \leqslant n$ now by definition $s(n) = n \bigcup \{n\} \Rightarrow n \in s(n) \Rightarrow n \leqslant s(n) \underset{4.46 \wedge 0 \leqslant n}{=} 0 \leqslant s(n) \Rightarrow s(n) \in S$

Using 4.10 we have then $S = \mathbb{N}_0$ proving our theorem. □

Theorem 4.48. *If* $n \in \mathbb{N}_0 \Rightarrow n < s(n)$

Proof. From $n \in n \bigcup \{n\} = s(n)$ we have $n \leqslant s(n) \underset{4.14}{\Rightarrow} n < s(n)$ □

Theorem 4.49. *If* $n \in \mathbb{N}_0$ *then* $k \in n \Leftrightarrow k < n$

Proof.

1. **($k \in n \Rightarrow k < n$)** We prove this by induction so let $S = \{n \in \mathbb{N}_0 | \text{if } k \in n \Rightarrow k < n\}$ then

   a. If $n = 0$ then $n = \emptyset \Rightarrow k \in n$ is impossible and thus '$k \in 0 \Rightarrow k < n$' is always true giving that $0 \in S$

   b. If $n \in S$ then if $k \in s(n) = n \bigcup \{n\}$ we have

      i. **($k \in n$)** then as $n \in S$ we have $k < n$ so $k \neq n$. From $k \in n \subseteq n \bigcup \{n\} = s(n)$ we have $k \leqslant s(n)$. Now if $k = s(n) = n \bigcup \{n\}$ then as $n \in \{n\} \Rightarrow n \in n \bigcup \{n\} \Rightarrow n \in k \Rightarrow n \leqslant k \underset{k < n \Rightarrow k \leqslant n \text{ and } 4.46}{\Rightarrow} n = k$ contradicting $k \neq n$ so that we have $k \neq s(n)$ and thus $k < s(n)$

      ii. **($k = n$)** then we have $k = n < s(n) \Rightarrow k < s(n)$

   from (i) and (ii) it follows then that $s(n) \in S$

   Using mathematical induction we have then $S = \mathbb{N}_0$ proving that if $n \in \mathbb{N}_0 = S$ then if $k \in n \Rightarrow k < n$

2. **($k < n \Rightarrow k \in n$)** If $k < n$ then $k \leqslant n$ and $k \neq n$ so $k \in n \vee k = n$ and $k \neq n$ giving $k \in n$ □

Theorem 4.50. *If* $k, n, m \in \mathbb{N}_0$ *then we have*

1. $k \leqslant n$ *and* $n < m$ *then* $k < m$
2. $k < n$ *and* $n \leqslant m$ *then* $k < m$

Proof. We have

1. If $k \leqslant n$ and $n < m$ then we have as $n < m \Rightarrow n \leqslant m$ that $k \leqslant m$. Now for k we have the following possibilities:

   a. **($k < n$)** then using the previous theorem we have $k \in n$ and $n \in m$. If now $k = m$ then we have $n \in k$. From $n \in k$ and $k \in n$ it follows then that $k \leqslant n$ and $n \leqslant k$ giving $k = n$ contradicting $k < n$. So we must have $k \neq m$ and thus $k < m$.

   b. **($k = n$)** if now $k = m$ then $m = n$ contradicting $n < m$ so we must have $k \neq m$ and thus we have $k < m$.

2. If $k<n$ and $n\leqslant m$ then we have $k\leqslant m$. Now for n we have the following possibilities

   a. $\boldsymbol{(n<m)}$ Using the previous theorem we have $n\in m$. If now $k=m$ then $n\in k\Rightarrow n\leqslant k$ this with $k<n\Rightarrow k\leqslant n$ means $k=n$ contradicting $k<n$. So we have $k\neq m$ and thus $k<m$.

   b. $\boldsymbol{(n=m)}$ then from $k<n \underset{\text{previous theorem}}{\Rightarrow} k\in n$ we have $k\in m \underset{\text{previous theorem}}{\Rightarrow} k<m$ □

Theorem 4.51. *For $n,m\in\mathbb{N}_0$ we have $n<m\Rightarrow s(n)\leqslant m$*

Proof. We prove this by mathematical induction, so let $S_n=\{m\in\mathbb{N}_0|n<m\Rightarrow s(n)\leqslant m\}$ we have then

1. As $n<0$ means that $n\in 0=\emptyset$ which is false so $n<0\Rightarrow s(n)\leqslant 0$ is true or $0\in S_n$

2. If $m\in S_n$ then $n<m\Rightarrow s(n)\leqslant m$, for $s(n)<m$ we have then $s(n)\in m$ we have now to prove that $s(m)\in S_n$ or

$$n<s(m) \;\;\Rightarrow\;\; s(n)\leqslant s(m)$$

   Now as $n<s(m)\Rightarrow n\in s(m)$ by 4.11 then we have either

   a. $n\in m\Rightarrow n<m\Rightarrow s(n)\leqslant m \underset{4.46 \text{ and } m\in s(m)=m\bigcup\{m\}\Rightarrow m\leqslant s(m)}{\Rightarrow} s(n)\leqslant s(m)$

   b. $n=m\Rightarrow s(n)=s(m)\Rightarrow s(n)\leqslant s(m)$

   we conclude thus that $s(m)\in S_n$

Using 4.10 we have then that $S_n=\mathbb{N}_0$. So if $n,m\in\mathbb{N}_0\Rightarrow m\in S_n\Rightarrow n<m\Rightarrow s(n)\leqslant m$ □

Theorem 4.52. *$\langle\mathbb{N}_0,\leqslant\rangle$ is well-ordered*

Proof. We prove this by contradiction, so let $\emptyset\neq A\subseteq\mathbb{N}_0$ be nonempty without a least element. Define then $S_A=\{n\in\mathbb{N}_0|\forall m\in A\,\text{we have}\,n\leqslant m\}$ then as A has no least element we must have $S_A\bigcap A=\emptyset$ [Otherwise $x\in S_A\bigcap A$ would mean that x is a least element of A].

We prove now by induction that $S_A=\mathbb{N}_0$

1. By 4.47 we have $\forall m\in A\subseteq\mathbb{N}_0$ that $0\leqslant m\Rightarrow 0\in S_A$

2. If $n\in S_A$ then $\forall m\in A\,\text{we have}\,n\leqslant m$. Now if $n\in A$ then n is a least element of A (which we assumed to be not the case) so we must have $\forall m\in A\vdash n<m$, using 4.51 we have then that $\forall m\in A\, s(n)\leqslant m$ and thus that $s(n)\in S_A$.

By 4.10 we have thus that $S_A=\mathbb{N}_0$ but this means that $S_A\bigcap A=\mathbb{N}_0\bigcap A=A\neq\emptyset$ contradicting that $S_A\bigcap A=\emptyset$. So every nonempty subset of $\mathbb{N}_0$ must have a least element. □

Theorem 4.53. *$\langle\mathbb{N}_0,\leqslant\rangle$ is fully-ordered*

Proof. This follows from the well-orderings of $\langle N, \leqslant\rangle$ and 2.183. $\square$

Theorem 4.54. $\langle \mathbb{N}_0, \leqslant\rangle$ *is conditional complete.*

Proof. This follows from the fact that $\langle \mathbb{N}_0, \leqslant\rangle$ is well-ordered and 2.184 $\square$

Theorem 4.55. $\forall n, m \in \mathbb{N}_0$ *then* $n < m \Leftrightarrow s(n) < s(m)$

Proof.

1. $\boldsymbol{(n < m)}$ From 4.51 we have then $s(n) \leqslant m$ and as $m < s(m)$ we have thus $s(n) < s(m)$
2. $\boldsymbol{(s(n) < s(m))}$ Assume that $m \leqslant n$ then we have either $m = n$ which leads to the contradiction $s(m) = s(n)$, or $m < n \underset{(1)}{\Rightarrow} s(m) < s(n) \Rightarrow s(m) < s(m)$ again a contradiction. So we must conclude that $n < m$. $\square$

Theorem 4.56. *If* $n, m, k \in \mathbb{N}_0$ *then* $n < m \Leftrightarrow n+k < m+k$. *Notice that from* $n = m \Leftrightarrow n+k = m+k$ *we have also the equivalence* $n \leqslant m \Leftrightarrow n+k \leqslant m+k$

Proof. We prove this by induction on k so let $B = \{k \in \mathbb{N}_0 | \text{if } n, m \in \mathbb{N}_0 \text{ then } n \leqslant m \Leftrightarrow n+k \Leftrightarrow m+k\}$ then we have

1. If $k = 0$ then clearly we have for $n, m \in \mathbb{N}_0$ that $n \leqslant m \Leftrightarrow n+0 \leqslant m+0 \Leftrightarrow n+k \Leftrightarrow m+k \Rightarrow 0 \in B$
2. If $k \in B$ then we have $n \leqslant m \Leftrightarrow n+k \Leftrightarrow m+k \underset{\text{previous theorem}}{\Leftrightarrow} s(n+k) = s(m+k) \underset{s(n+k) = n+s(k), s(m+k) = m+s(k)}{\Leftrightarrow} n+s(k) = m+s(k)$ so we have $s(k) \in B$

using mathematical induction we have $B = \mathbb{N}_0$ proving our theorem. $\square$

Theorem 4.57. *If* $n, m \in \mathbb{N}_0$ *with* $n < s(m)$ *then* $n \leqslant m$

Proof. If we would assume the opposite then $m < n$ and using (4.51) we have then $s(m) \leqslant n < s(m)$ a contradiction. So we must have $n \leqslant m$ $\square$

Theorem 4.58. *If* $n, m \in \mathbb{N}_0$ *and* $n < m$ *then there exists a* $k \in \mathbb{N}_0, k \neq 0$ *such that* $n+k = m$

Proof. Let $S_n = \{m \in \mathbb{N}_0 | n < m \Rightarrow \exists k \in \mathbb{N}_0 | k \neq 0 \vdash n+k = m\}$ then we have

1. $n < 0$ is false as from 4.47 $0 \leqslant n \Rightarrow 0 < 0 \Rightarrow 0 \neq 0$ a contradiction so we have that $n < 0 \Rightarrow \exists k \in \mathbb{N}_0, k \neq 0 \vdash n + k = m$ is true, so $0 \in S_n$
2. If $m \in S_n$ suppose now that $n < s(m)$ then as $\mathbb{N}_0$ is fully ordered (see 4.53) we have the following cases
   a. $m \leqslant n$ here we have the following cases
      i. $m < n \underset{4.51}{\Rightarrow} s(m) \leqslant n \Rightarrow s(m) < s(m)$ a contradiction so this case does not apply.
      ii. $m = n \Rightarrow s(n) = s(m)$ contradiction $n < s(m)$ so this case does not apply.

b. $n \leqslant m$ here we have the following possibilities

i. $n = m$ then $n+1 \underset{4.30}{=} s(n) = s(m) \underset{1 \neq 0}{\Rightarrow} s(m) \in S_n$

ii. $n < m \underset{m \in S_s}{\Rightarrow} \exists k \in \mathbb{N}_0, k \neq 0 \vdash n+k = m \Rightarrow n+s(k) \underset{4.31}{=} s(n+k) = s(m)$. As we have $k \in s(k) = k \bigcup \{k\} \Rightarrow k \leqslant s(k)$ and $0 < k \Rightarrow 0 < s(k) \Rightarrow 0 \neq s(k)$ we conclude that $s(m) \in S_n$

Using mathematical induction 4.10 we have then that $S_n = \mathbb{N}_0$. So if $n, m \in \mathbb{N}_0 \Rightarrow m \in S_n \Rightarrow$ if $n < m \Rightarrow \exists k \in \mathbb{N}_0, k \neq 0 \vdash n + k = m$. $\square$

Lemma 4.59. *If $n \in \mathbb{N}_0,\ k \in \mathbb{N} \Rightarrow n < n+k$*

Proof. We prove this by mathematical induction, so let $k \in \mathbb{N}$ and $S_k = \{n \in \mathbb{N}_0 | n < n+k\}$ then we have

1. $0+k \underset{4.29}{=} k$ and as $k \neq 0$ we have $0 < k = 0+k \Rightarrow 0 \in S_k$
2. If $n \in S_k$ then by 4.48 we have $n+k < s(n+k) \underset{4.31}{=} s(n)+k$ also as $n \in S_k$ we have $n < n+k \Rightarrow s(n) \leqslant n+k \Rightarrow s(n) < s(n)+k \Rightarrow s(n) \in S_k$

By 4.10 we have $S_k = \mathbb{N}_0$. So if $k \in \mathbb{N}$ and $n \in \mathbb{N}_0 \Rightarrow n \in S_k$ and thus $n < n+k$ $\square$

Theorem 4.60. *If $n, k \in \mathbb{N}_0$ then $n+k = 0 \Rightarrow n = k = 0$*

Proof. Suppose that $k \neq 0$ then as $0 \leqslant n \underset{4.59}{\Rightarrow} 0 \leqslant n < n+k = 0 \Rightarrow 0 < 0 \Rightarrow 0 \neq 0$ and thus we reach a contradiction. So $k = 0$ but then $n = n+0 = n+k = 0$ and thus also $n = 0$ $\square$

Theorem 4.61. *If $n, m \in \mathbb{N}_0$ then $n < m \Leftrightarrow \exists! k \in \mathbb{N} \vdash n+k = m$.*

Proof.

$\Rightarrow$

We prove this by mathematical induction, so let $S_n = \{m \in \mathbb{N}_0 | \text{if } n < m \Rightarrow \exists k \in \mathbb{N} \vdash n+k = m\}$ then we have

1. We have by 4.47 that $0 \leqslant n$ so if $n < 0 \Rightarrow 0 < 0$ a contradiction. So we can never have $n < 0 \Rightarrow 0 \in S_n$
2. If $m \in S_n$ suppose then that $n < s(m)$ then we have the following possible cases for m

   a. $n = m \Rightarrow s(m) = s(n) \underset{4.30}{=} n+1 \underset{1 \neq 0}{\Rightarrow} s(m) \in S_n$

   b. $n < m \underset{m \in S_n}{\Rightarrow} \exists k \in \mathbb{N} \vdash n+k = m \underset{4.31}{\Rightarrow} n+s(k) = s(n+k) = s(m)$ and as from $0 \leqslant k \underset{k \neq 0}{\Rightarrow} 0 < k$ and by 4.48 $k < s(k) \Rightarrow 0 < s(k) \Rightarrow s(k) \neq 0$ and as we just have proved that $n+s(k) = s(m)$ we have $s(m) \in S_n$

   c. $m < n \underset{4.51}{\Rightarrow} s(m) \leqslant n \underset{n < s(m)}{\Rightarrow} s(m) < s(m)$ is a contradiction so this case does not apply

   In all the cases that applies we had $s(m) \in S_n$

Using 4.10 we have then $S_n = \mathbb{N}_0$ so if $m, n \in \mathbb{N}_0 \Rightarrow m \in S_n \Rightarrow$ if $n < m \Rightarrow \exists k \in \mathbb{N} \vdash n+k = m$. Finally to prove uniqueness assume that $n+k = m = n+k'$ then by 4.44 we have $k = k'$.

$\Leftarrow$

Suppose $\exists k \in \mathbb{N} \vdash n+k=m$ then for n, m we have by 4.53 that either $n=m$, $m<n$ or $n<m$. Now if $n=m$ then as $k \in \mathbb{N}$ we have by 4.59 that $n<n+k=m=n$ a contradiction. Also if $m<n$ then again from 4.59 we have $n<n+k=m \Rightarrow n<n$ again a contradiction. So the only possibility left is $n<m$.

□

Corollary 4.62. *If $n, m \in \mathbb{N}_0$ then $n \leqslant m \Leftrightarrow \exists! k \in \mathbb{N}_0 \vdash n+k=m$.*

Proof.

$\Rightarrow$ if $n \leqslant m$ then we have either

1. $n<m \underset{\text{previous theorem}}{\Rightarrow} \exists! k \in \mathbb{N} \vdash n+k=m \Rightarrow \exists k \in \mathbb{N}_0 \vdash n+k=m$
2. $n=m \Rightarrow m=n+0 \Rightarrow \exists k \in \mathbb{N}_0\, (k=0) \vdash n+k=m$

Finally to prove uniqueness assume that $n+k=m=n+k'$ then by 4.44 we have $k=k'$.

$\Leftarrow$ If $\exists k \in \mathbb{N}_0 \vdash n+k=m$ then we have for k the following possibilities

1. $k=0 \Rightarrow n=m \Rightarrow n \leqslant m$
2. $k \neq 0 \Rightarrow k \in \mathbb{N}$ and using the previous theorem we have then $n<m \Rightarrow n \leqslant m$ □

Definition 4.63. *If $n, m \in \mathbb{N}_0$, $n \leqslant m$ then the unique k that by the previous theorem exists so that $n+k=m$ is noted by $m-n$ so $n+(m-n)=m$.*

***Note* 4.64.** *If $n \in \mathbb{N}_0 \Rightarrow n-n=0$ (this is trivial because of uniqueness and the fact that $n+0=n$*

Theorem 4.65. *If $n, m \in \mathbb{N}_0$ and $i \in \mathbb{N}_0$ such that $n \leqslant i \leqslant m$ then $0 \leqslant i-n \leqslant m-n$*

Proof. First if $m-n<i-n \Rightarrow m=(m-n)+n<(i-n)+n=i \Rightarrow m<i \leqslant m \Rightarrow m<m$ a contradiction so we have that $i-n \leqslant m-n$. By applying the above on $n \leqslant i$ we find that $0=n-n \leqslant i-n \Rightarrow 0 \leqslant i-n$ □

Theorem 4.66. *If $n \in \mathbb{N}_0$ then there does not exists a $k \in \mathbb{N}_0$ such that $n<k<s(n)$*

Proof. We prove this by mathematical induction, so take $S=\{n \in \mathbb{N}_0 |$ there does not exists $k \in \mathbb{N}_0$ such that $n<k<s(n)\}$. We have then

1. If $0<k<1 \underset{4.51}{\Rightarrow} s(0)=1 \leqslant s(k) \leqslant 1 \Rightarrow s(k)=1=s(0) \underset{4.15}{\Rightarrow} k=0$ contradicting $0<k$. So we conclude that $0 \in S$
2. Let $n \in S$. Assume now that $\exists k \in \mathbb{N}_0$ such that $s(n)<k<s(s(n))$ then from 4.61 we have $\exists l \in \mathbb{N}$ such that $k=s(n)+l=s\,(n+l)$. Take then $k'=n+l$ then as $l \in \mathbb{N}$ we have by 4.61 again that $n<k'$. From $s(k')=k<s(s(n)) \Rightarrow s(k')<s(s(n)) \underset{4.55}{\Rightarrow} k'<s(n)$ so we have $n<k'<s(n)$ contradicting $n \in S$. So the assumption turns out to be wrong and thus $s(n) \in \mathbb{N}_0$

So by mathematical induction we have $S=\mathbb{N}_0$ proving our theorem. □

Theorem 4.67. *If* $\emptyset \neq A \subseteq \mathbb{N}_0$ *is a set such that* $\sup(A)$ *exists then* $\sup(A) \in A$

Proof. Suppose that $A \neq \emptyset$ and that $\sup(A)$ exists. We have now the following cases for $\sup(A)$

1. **($\sup(A) = 0$)** As $A \neq \emptyset$ there exists a $x \in A$ and as a sup is a upper bound we have $x \leqslant 0$ and from 4.47 we have $0 \leqslant x \Rightarrow x = 0 \Rightarrow \sup(A) \in A$
2. **($\sup(A) \neq 0$)** Using 4.17 we have $\exists n \in \mathbb{N}_0$ such that $\sup(A) = s(n)$. Then as $n < s(n) = \sup(A)$ there must exists a $x \in A$ with $n < x \leqslant \sup(A)$ (otherwise n would be the lowest upper bound). If $x \neq \sup(A)$ then we have $n < x < s(n)$ which is forbidden by 4.66 so we must have $x = \sup(A) \Rightarrow \sup(A) \in A$. □

Theorem 4.68. *If* $n, m, r, s \in \mathbb{N}_0$ *then if* $n < m \wedge r < s \Rightarrow n+r < m+s$

Proof. If $n < m$ and $r < s$ then there exists $k, l \in \mathbb{N}$ such that $m = n+k$ and $s = r+l$ so $m+s = (n+k)+(r+l) = (n+r)+(k+l)$. Now as $0 < k$ we have by 4.59 that $0 < k+l \Rightarrow k+l \in \mathbb{N}$. Using 4.61 we have then that $n+r < m+s$. □

Lemma 4.69. *If* $n, m \in \mathbb{N}$ *then* $n \cdot m \in \mathbb{N}$

Proof. We prove this by mathematical induction. If $n \in \mathbb{N}$ then let $S_n = \{m \in \mathbb{N}_0 | m \neq 0 \Rightarrow n \cdot m \neq 0\}$ then we have

1. $0 \in S$ [$0 \neq 0$ is false so [$0 \neq 0 \Rightarrow n \cdot m \neq 0$] is true]
2. If $m \in S_n$ then $n \cdot s(m) \underset{4.39}{=} n \cdot m + n$ as $m \in S_n$ we have $n \cdot m \in S_n \Rightarrow 0 < n \cdot m \underset{4.59 \text{ and } n \in S_n}{\Rightarrow} 0 < n \cdot m < n \cdot m + n \Rightarrow 0 < n \cdot m + n = n \cdot s(m) \Rightarrow n \cdot s(m) \neq 0$

Using 4.10 we have then that $S_n = \mathbb{N}_0$ So if $n, m \in \mathbb{N} \Rightarrow 0 \neq m \in S_n \Rightarrow n \cdot m \neq 0$ □

Theorem 4.70. *If* $n, m, k \in \mathbb{N}_0$ *then if* $k \neq 0$ *and* $n < m \Rightarrow k \cdot n < k \cdot m$

Proof. As $n < m \underset{4.61}{\Rightarrow} \exists l \in \mathbb{N}$ such that $n+l = m \Rightarrow k \cdot (n+l) = k \cdot m \Leftrightarrow k \cdot n + k \cdot l = k \cdot m$. As we have $0 \leqslant n$ and $n < m \Rightarrow 0 < m \Rightarrow m \in \mathbb{N} \underset{4.69}{\Rightarrow} k \cdot l \in \mathbb{N} \underset{4.61}{\Rightarrow} k \cdot n < k \cdot m$ □

Corollary 4.71. *If* $n, m, k \in \mathbb{N}_0$ *then if* $n \leqslant m \Rightarrow n \cdot k \leqslant m \cdot k$

Proof. If $n \leqslant m$ then we have the following cases

1. $n = m \Rightarrow n \cdot k = m \cdot k \Rightarrow n \cdot k \leqslant m \cdot k$
2. $n < m$ then we have for k the following cases
   a. $k = 0 \underset{4.37}{\Rightarrow} n \cdot k = 0 = m \cdot k \Rightarrow n \cdot k \leqslant m \cdot k$
   b. $k \neq 0 \underset{\text{previous theorem}}{\Rightarrow} n \cdot k < m \cdot k \Rightarrow n \cdot k \leqslant m \cdot k$ □

Theorem 4.72. *If* $n, m \in \mathbb{N}_0, k \in \mathbb{N}$ *and* $n \cdot k = m \cdot k \Rightarrow n = m$

Proof. We have by 4.53 that $n < m$,$m < n$ or $n = m$. Now if $n < m \underset{4.70}{\Rightarrow} n \cdot k < m \cdot k$ a contradiction, also if $m < n \Rightarrow m \cdot k < n \cdot k$ a contradiction. So the only thing left is that $n = m$. □

Theorem 4.73. (Archimedean property of $\mathbb{N}_0$) *If $x, y \in \mathbb{N}_0$ and $x > 0$ then there exists a $z \in \mathbb{N}_0$ such that $z \cdot x > y$*

Proof. We have the following possibilities for y

1. $y \leqslant x$ then as $1 < 2 = s(1)$ (see 4.48) we have by 4.70 that $y \leqslant x = x \cdot 1 < 2 \cdot x$ so taking $z = 2$ we have $z \cdot z > y$
2. $x < y$ then by 4.61 there exists a $k \in \mathbb{N}$ such that $y = x + k$ now as $0 < x$ we have by 4.51 that $1 = s(0) \leqslant x \Rightarrow 1 \leqslant x$, also we have that $x + k < s(x + k)$ This gives using 4.71 $s(x+k) \cdot 1 \leqslant s(x+k) \cdot x \Rightarrow y = x + k < s(x+k) \leqslant s(x+k) \cdot x \Rightarrow y < s(x+k) \cdot y$, so taking $z = s(x+k)$ we have $z \cdot x > y$ □

Theorem 4.74. (Division Algorithm) *If $m, n \in \mathbb{N}_0$ and $n > 0$ then there exists a unique $r \in \mathbb{N}_0$ with $0 \leqslant r < n$ and a unique $q \in \mathbb{N}_0$ such that $m = n \cdot q + r$*

Proof. First we prove the existence of $r, q \in \mathbb{N}_0$. For m we have the following cases to consider

1. $(\boldsymbol{m = 0})$ In this case we can take $r = 0$, $q = 0$ and then $0 = r < n$ and $n \cdot q + r = n \cdot 0 + 0 = 0 = m$
2. $(\boldsymbol{0 < m})$ We have then the following cases for n
   a. $(\boldsymbol{n = 1})$ Then we can take $q = m, r = 0$ and then $0 = r < n$ and $n \cdot q + r = 1 \cdot m + 0 = m$
   b. $(\boldsymbol{1 < n})$ Then from 4.70 we have $m < n \cdot m$ so $m \in A_{n,m} = \{x \in \mathbb{N}_0 | m < n \cdot x \wedge x \leqslant m\} \Rightarrow A_{n,m} \neq \emptyset$. Using well-ordering (see 4.52) there exists a $q' = \min(A_{n,m})$. We have $q' \neq 0$ [because if $q' = 0$ then from $q' \in A_{n,m}$ we have $0 < m < n \cdot 0 = 0 \Rightarrow 0 < 0$ a contradiction] and thus there exists a $q \in \mathbb{N}_0$ such that $q' = s(q) = q + 1$. From $q < q' \leqslant m$ we have by the definition of the minimum that $q \notin A_{n,m}$ and thus $n \cdot q \leqslant m$ so we can define $r = m - n \cdot q$ where $r \in \mathbb{N}_0$ and thus $\boldsymbol{m = n \cdot q + r}$. If now $n \leqslant r \Rightarrow n + n \cdot q \leqslant r + n \cdot q = m \Rightarrow n \cdot (q + 1) \leqslant m \Rightarrow n \cdot q' \leqslant m$ and from $q' \in A_{n,m}$ we have $m < n \cdot q'$ so we reach the contradiction $n \cdot q' \leqslant m < n \cdot q' \Rightarrow n \cdot q < n \cdot q'$ and we conclude that $\boldsymbol{0 \leqslant r < n}$.

Now to prove uniqueness assume that $m = n \cdot q + r = n \cdot q'' + r''$ then we have if $q \neq q''$ that either the following is true

1. $(\boldsymbol{q < q''})$ then $q + 1 \leqslant q'' \Rightarrow n \cdot (q+1) \leqslant n \cdot q'' \Rightarrow n \cdot q + n \leqslant n \cdot q'' \Rightarrow n \cdot q + n + r + r'' \leqslant n \cdot q'' + r + r'' \Rightarrow m + n + r'' \leqslant m + r \Rightarrow n + r'' \leqslant r \Rightarrow n \leqslant r$ contradicting $r < n$
2. $(\boldsymbol{q'' < q})$ then $q'' + 1 < q \Rightarrow n \cdot (q''+1) \leqslant n \cdot q \Rightarrow n \cdot q'' + n \leqslant n \cdot q \Rightarrow n \cdot q'' + n + r + r'' \leqslant n \cdot q + r + r'' \Rightarrow m + n + r \leqslant m + r'' \Rightarrow n + r \leqslant r'' \Rightarrow n \leqslant r''$ contradicting $r'' < n$

as neither of the above cases can be true we must have $q = q''$ but then $n \cdot q = n \cdot q'' \underset{n \cdot q + r = n \cdot q'' + r''}{\Rightarrow} r = r''$ proving uniqueness. □

4.4.1 Other forms of Mathematical induction

Once we have discussed the order relation on $\mathbb{N}_0$ we can write down some other forms of mathematical induction.

Definition 4.75. Given $n\in\mathbb{N}_0$ *we define* $\{n,\ldots\}$ *to be equal to* $\{i\in\mathbb{N}_0|n\leqslant i\}$

Definition 4.76. *Given* $n,m\in\mathbb{N}_0$ *with* $n\leqslant m$ *we define* $\{n,\ldots,m\}=\{i\in\mathbb{N}_0|n\leqslant i\wedge i\leqslant m\}$, *note that if* $m<n$ *then* $\{n,\ldots,m\}=\emptyset$

We use this in the following theorem

Theorem 4.77. *If* $n\in\mathbb{N}_0$ *and* $X\subseteq\{n,\ldots\}$ *is such that*

1. $n\in X$
2. *if* $i\in X\Rightarrow i+1\in X$

then $X=\{n,\ldots\}$

Proof. Take $B=\{i\in\mathbb{N}_0|i+n\in X\}$ then we have

1. As $n\in X$ we have $0+n\in X\Rightarrow 0\in B$
2. If $i\in B$ then $i+n\in X\Rightarrow(i+1)+n=(i+n)+1\in X\Rightarrow i+1\in B$

by mathematical induction (see 4.10) we have then $B=\mathbb{N}_0$. If now $i\in\{n,\ldots\}$ then $n\leqslant i\Rightarrow i-n\in\mathbb{N}_0=B\Rightarrow(i-n)+n\in X\Rightarrow i\in X\Rightarrow\{n,\ldots\}\subseteq X\subseteq\{n,\ldots\}\Rightarrow X=\{n,\ldots\}$ □

Theorem 4.78. *If* $n\in\mathbb{N}_0$ *and* $P(i)$ *is a predicate defined* $\forall i\in\{n,\ldots\}$ *depending on* i *is such that*

1. $P(n)$ *is true*
2. *If* $n\leqslant i$ *and* $P(i)$ *is true then* $P(i+1)$ *is true*

then $\forall i\in\{n,\ldots\}$ *we have that* $P(i)$ *is true*

Proof. Define $X=\{i\in\{n,\ldots\}|P(i)\,\text{is true}\}$ then we have

1. $n\in X$
2. if $i\in X\Rightarrow i+1\in X$

thus $X=\{n,\ldots\}$ and if $i\in\{n,\ldots\}\Rightarrow i\in X\Rightarrow P(i)$ is true □

The last form is the form of induction used in most mathematical text. In this text however we use mostly the form 4.77 as I think this forces you to express the induction hypotheses in a much clearer form. As a example of induction and recursion lets define the composition of n functions.

Definition 4.79. *Let* $n\in\mathbb{N}$, $\{X_i\}_{i\in\{1,\ldots,n+1\}}$ *a family of sets,* $\{f_i\}_{i\in\{1,\ldots,n\}}$ *with* $\forall i\in\{1,\ldots,n\}$ *we have* $f_i\subseteq X_i\times X_{i+1}$ *then* $f_n\circ\ldots\circ f_1\subseteq X_1\times X_{n+1}$ *is defined by*

1. $f_1\circ\ldots\circ f_1=f_1$
2. $f_{n+1}\circ\ldots.\circ f_1=f_{n+1}\circ(f_n\circ\ldots\circ f_1)$

Theorem 4.80. *Let $n \in \mathbb{N}$, $\{X_i\}_{i \in \{1,\ldots,n+1\}}$ a family of sets, $\{f_i\}_{i \in \{1,\ldots,n\}}$ a family of functions $f_i: X_i \to X_{i+1}$ then $f_n \circ \ldots \circ f_1: X_1 \to X_{n+1}$ is a function*

Proof. We use mathematical induction to prove this so let $S = \{n \in \mathbb{N} | \text{if } \{f_i\}_{i \in \{1,\ldots,n\}} \text{ is } a \text{ family of functions } f_i: X_i \to X_{i+1} \text{ then } f_n \circ \ldots. \circ f_1: X_1 \to X_{n+1} \text{ is } a \text{ function}\}$, then we have

$1 \in S$. If $\{f_i\}_{i \in \{1,\ldots,1\}}$ is a family of functions $f_i: X_i \to X_{i+1}$ then $f_1: X_1 \to X_2$ is a function so that $f_1 \circ \ldots \circ f_1 = f_1: X_1 \to X_2$ is a function proving that $1 \in S$

$n \in S \Rightarrow n+1 \in S$. If $\{f_i\}_{i \in \{1,\ldots,n+1\}}$ is a family of functions $f_i: X_i \to X_{i+1}$ then as $n \in S$ we have $f_n \circ \ldots \circ f_1: X_1 \to X_{n+1}$ is a function, then as $f_{n+1}: X_{n+1} \to X_{n+2} = X_{(n+1)+1}$ is a function by the hypothesis we have that $f_{n+1} \circ \ldots \circ f_1 = f_{n+1} \circ (f_n \circ \ldots \circ f_1): X_1 \to X_{(n+1)+1}$ is a function proving that $n+1 \in S$

Using mathematical induction we have that $S = \mathbb{N}$ proving the theorem. □

We use now the associativity of the composition of functions to prove a more general result

Theorem 4.81. *Let $n \in \mathbb{N} \setminus \{1\}$, $m \in \mathbb{N}_0$ $\{X_i\}_{i \in \{1,\ldots,n+1\}}$ a family of sets, $\{f_i\}_{i \in \{1,\ldots,n+m\}}$ a family of functions $f_i: X_i \to X_{i+1}$ then we have that $(f_{n+m} \circ \ldots \circ f_n) \circ (f_{n-1} \circ \ldots \circ f_1) = (f_{n+m} \circ \ldots \circ f_1)$*

Proof. We prove this by induction on m so let $S_n = \{m \in \mathbb{N}_0 | (f_{n+m} \circ \ldots \circ f_n) \circ (f_{n-1} \circ \ldots. \circ f_1)\}$ then we have

$0 \in S_n$. In this case we have $(f_{n+0} \circ \ldots \circ f_n) \circ (f_{n-1} \circ \ldots \circ f_1) = (f_n \circ \ldots \circ f_n) \circ (f_{n-1} \circ \ldots \circ f_1) = f_n \circ (f_{n-1} \circ \ldots \circ f_1) = (f_n \circ \ldots \circ f_1)$ proving that $0 \in S$

$m \in S_n \Rightarrow m+1 \in S_n$. Then for $(f_{n+(m+1)} \ldots f_n) \circ (f_{n-1} \circ \ldots \circ f_1) = ((f_{n+m} \circ (f_{n+m} \circ \ldots \circ f_n)) \circ (f_{n-1} \odot . \circ f_1)) \underset{2.34}{=} f_{n+(m+1)} \circ ((f_{n+m} \circ \ldots \circ f_n) \circ (f_{n-1} \circ \ldots \circ f_1)) \underset{m \in S_n}{=} f_{n+(m+1)} \circ (f_{n+m} \circ \ldots \circ f_1) = (f_{n+(m+1)} \circ \ldots \circ f_1)$ proving that $m+1 \in S$

Mathematical induction proves then the theorem. □

Chapter 5
Finite and Infinite Sets

5.1 Introduction

Definition 5.1. *Two classes A and B are called equipotent if there exists a bijection $f: A \rightarrow B$, we note this as $A \approx B$*

Definition 5.2. *Given two classes A and B then $A \preccurlyeq B$ if there is a $C \subseteq B$ such that $A \approx C$*

Definition 5.3. *Given two classes A and B then $A \prec B$ if $A \preccurlyeq B$ and $\neg(A \approx B)$*

Theorem 5.4. *Given two classes A and B then $A \preccurlyeq B$ if and only if there exists a injection $f: A \rightarrow C$ where $C \subseteq B$*

Proof.

$$\begin{array}{rcl} A \preccurlyeq B & \Leftrightarrow & \text{there exists a } C \subseteq B \text{ and a bijection } f: A \rightarrow C \\ & \Rightarrow & f: A \rightarrow B \text{ is a injection} \\ f: A \rightarrow B \text{ is a injection} & \Rightarrow & f: A \rightarrow f(A) \text{ is a bijection} \\ & \underset{f(A) \subseteq B}{\Rightarrow} & A \preccurlyeq B \end{array}$$

□

Theorem 5.5. *If A is a set then there exists no surjective function between A and $\mathcal{P}(A)$*

Proof. We prove this by contradiction so assume that there is a surjective function $f: A \rightarrow \mathcal{P}(A)$ define then $B = \{x \in A | x \notin f(x)\}$. As $B \subseteq A$ we have that $B \in \mathcal{P}(A)$ and by surjectivity of f there exists a $y \in A$ such that $B = f(y)$. Now if $y \in B$ then $y \notin f(y) = B \Rightarrow y \notin B$ a contradiction. If $y \notin B$ then as $B = f(y)$ we have $y \notin f(y)$ and thus $y \in B$ again a contradiction. So all the cases $y \in B$ or $y \notin B$ gives a contradiction and thus a surjective $f: A \rightarrow \mathcal{P}(A)$ can not exist. □

Corollary 5.6. *If A is a set then no subset of A can be equipotent with $\mathcal{P}(A)$ (or with 2^A)*

Proof. First we prove that no subset of A can be equipotent with $\mathcal{P}(A)$. If $B \subseteq A$ then we have the following possible cases:

1. **($\boldsymbol{B = A}$)** then by 5.5 we can not have a bijection (which is surjective) between A and $\mathcal{P}(A)$, so B can not be equipotent with $\mathcal{P}(A)$

2. **($B \subset A$)** in this case $A \backslash B$ is non-empty and $B \bigcap (A \backslash B) = \emptyset$. Assume now that B is equipotent with $\mathcal{P}(A)$ then a bijection $f' : B \rightarrow \mathcal{P}(A)$ exists. Form then the function (see 2.31) $f : A \rightarrow \mathcal{P}(A)$ defined by $x \rightarrow f(x) = \begin{cases} \emptyset \text{ if } x \in A \backslash B \\ f'(x) \text{ if } x \in B \end{cases}$. If now $y \in \mathcal{P}(A)$ there exists by bijectivity of f' a $x \in B$ such that $f'(x) = y \Rightarrow f(x) = y \Rightarrow f$ is surjective which is impossible by 5.5. So we must have that B is not equipotent with $\mathcal{P}(A)$

Next if $B \subseteq A$ and B is equipotent with 2^A then there exists a bijection $f : B \rightarrow 2^A$. By 2.73 2^A and $\mathcal{P}(A)$ are bijective so there exists a bijection $g : 2^A \rightarrow \mathcal{P}(A)$, but this means that $g \circ f : B \rightarrow \mathcal{P}(A)$ is bijective which we have just proved is impossible. So B and $\mathcal{P}(A)$ cannot be equipotent. □

Corollary 5.7. *If A is a set then A cannot be equipotent with any $B \supseteq \mathcal{P}(A)$ or with any $B \supseteq 2^A$*

Proof. Suppose that A is equipotent with $B \supseteq \mathcal{P}(A)$ then there exists a bijection $f : A \rightarrow B$, using 2.61 we have that ${}_{B|}f : f^{-1}(\mathcal{P}(A)) \rightarrow \mathcal{P}(A)$ is a bijection which by the previous corollary is impossible.

Suppose that A is equipotent with $B \supseteq 2^A$ then there exists a bijection $f : A \rightarrow B$, using 2.61 we have that ${}_{B|}f : f^{-1}(2^A) \rightarrow 2^A$ is a bijection which by the previous corollary is impossible. □

Theorem 5.8. *If A, B are classes then there exists a injection $f : A \rightarrow B$ if and only if there exists a surjection $g : B \rightarrow A$*

Proof.

If there exists a injection $f : A \rightarrow B$ then by 2.44 there exists a function $g : B \rightarrow A$ such that $g \circ f = i_A$. Then if $y \in A$ we have $y = i_A(y) = g(f(y)) \Rightarrow g$ is surjective.

If there exists a surjection $g : B \rightarrow A$ then by 2.45 there exists a $f : A \rightarrow B$ such that $g \circ f = i_A$. Given $x, x' \in A$ if $f(x) = f(x') \Rightarrow g(f(x)) = g(f(x')) \Rightarrow (g \circ f)(x) = (g \circ f)(x') = i_A(x) = i_A(x') \Rightarrow x = x'$ and this we have that f is injective. □

Using the above theorem and 5.4 we have the following corollary

Corollary 5.9. *If A, B are classes then $A \preccurlyeq B$ if and only if there exists a surjection $f : B \rightarrow A$*

Theorem 5.10. *If A, B, C, D are classes where $A \bigcap C = \emptyset$ and $B \bigcap D$ are sets and $f : A \rightarrow B$ and $g : C \rightarrow D$ are bijections then $g \bigcup f : A \bigcup C \rightarrow B \bigcup D$ is a bijection*

Proof. As $f : A \rightarrow B$ and $g : C \rightarrow D$ are functions then clearly $f : A \rightarrow B \bigcup D$ and $g : C \rightarrow B \bigcup D$ are also functions and thus by 2.31 $f \bigcup g : A \bigcup C \rightarrow B \bigcup D$ is a function. Now as f, g are bijective we have the existence of the bijections $f^{-1} : B \rightarrow A$, $g^{-1} : D \rightarrow C$ meaning that $f^{-1} : B \rightarrow A \bigcup C$ and $g^{-1} : D \rightarrow A \bigcup C$ are functions and by 2.31 we have $f^{-1} \bigcup g^{-1} : B \bigcup D \rightarrow A \bigcup C$ is also a function. Now $\forall x \in B \bigcup D$ we have either

1. **($x \in B$)** then $(f^{-1} \bigcup g^{-1})(x) = f^{-1}(x) \in A \Rightarrow (f \bigcup g)((f^{-1} \bigcup g^{-1})(x)) = (f \bigcup g)(f^{-1}(x)) = f(f^{-1}(x)) = i_B(x) = x = i_{B \bigcup D}(x)$

2. **($x \in D$)** then $(f^{-1} \bigcup g^{-1})(x) = g^{-1}(x) \in C \Rightarrow (f \bigcup g)((f^{-1} \bigcup g^{-1})(x)) = (f \bigcup g)(g^{-1}(x)) = g(g^{-1}(x)) = i_D(x) = x = i_{B \bigcup D}(x)$

So we have $(f \bigcup g) \circ (f^{-1} \bigcup g^{-1}) = i_{B \bigcup D}$. Also $\forall y \in A \bigcup C$ we have either

1. **($y \in A$)** then $(f \bigcup g)(y) = f(y) \in B$ so that $(f^{-1} \bigcup g^{-1})((f \bigcup g)(y)) = (f^{-1} \bigcup g^{-1})(f(y)) = f^{-1}(f(y)) = i_A(y) = y = i_{A \bigcup C}(y)$
2. **($y \in C$)** then $(f \bigcup g)(y) = g(y) \in D$ so that $(f^{-1} \bigcup g^{-1})((f \bigcup g)(y)) = (f^{-1} \bigcup g^{-1})(g(y)) = g^{-1}(g(y)) = i_D(y) = y = i_{A \bigcup C}(y)$

so we have then $(f^{-1} \bigcup g^{-1}) \circ (f \bigcup g) = i_{A \bigcup C}$. This proves that $f \bigcup g: A \bigcup C \rightarrow B \bigcup D$ is a bijective function with inverse function $f^{-1} \bigcup g^{-1}: B \bigcup D \rightarrow A \bigcup C$ $\square$

Corollary 5.11. *If A, B, C, D are classes with $A \bigcap C = \emptyset$, $B \bigcap D = \emptyset$ and $A \approx B$ and $C \approx D$ then $(A \bigcup C) \approx (B \bigcup D)$*

Proof. This is ease as $A \approx B$, $C \approx D$ implies that there exists bijections $f: A \rightarrow B$ and $g: C \rightarrow D$. Using the previous theorem we haven the bijection $f \bigcup g: A \bigcup C \rightarrow B \bigcup D$ and thus that $(A \bigcup C) \approx (B \bigcup D)$ $\square$

Theorem 5.12. *If A, B, C, D are classes and $A \approx C$ and $B \approx D$ then $(A \times B) \approx (C \times D)$*

Proof. As $A \approx C$ and $B \approx D$ we have the existence of the bijections $f: A \rightarrow C$ and $g: B \rightarrow D$. Define $h: A \times B \rightarrow C \times D$ by $(x, y) \rightarrow (f(x), g(y))$ then we have

1. If $(x, y), (x', y') \in A \times B$ such that $h((x, y)) = h((x', y'))$ then $(f(x), g(y)) = (f(x'), g(y')) \Rightarrow f(x) = f(x') \wedge g(y) = g(y') \Rightarrow x = x' \wedge y = y' \Rightarrow (x, y) = (x', y')$ proving injectivity.
2. If $(u, v) \in C \times D$ then $u \in C \wedge v \in D \underset{f, g \, \text{are} \, \text{surjective}}{\Rightarrow} \exists x \in A, y \in B \vdash u = f(x) \wedge v = g(y) \Rightarrow (u, v) = (f(x), g(y)) = h((x, y))$

This proves that $h: A \times B \rightarrow C \times D$ is bijective and thus $(A \times B) \approx (C \times D)$ $\square$

Theorem 5.13. *If A, B, C, D are sets such that $A \approx B$ and $C \approx D$ then $A^C \approx B^D$*

Proof. As $A \approx B$ and $C \approx D$ there exists bijections $f: A \rightarrow B$ and $g: D \rightarrow C$. Define now $h: A^C \rightarrow B^D$ by $s \in A^C \rightarrow h(s) = f \circ s \circ g \in B^D$ [as $s: C \rightarrow A$ we have that $f \circ s \circ g: D \rightarrow B \Rightarrow f \circ s \circ g \in B^D$]. We have now

1. If $s, t \in A^C$ is such that $h(s) = h(t)$ then $f \circ s \circ g = f \circ t \circ g \Rightarrow f^{-1} \circ (f \circ s \circ g) = f^{-1} \circ (f \circ t \circ g) \Rightarrow s \circ g = t \circ g \Rightarrow (s \circ g) \circ g^{-1} = (t \circ g) \circ g^{-1} = s = t$ proving injectivity.
2. If $u \in B^D$ then $u: D \rightarrow B$ and thus $f^{-1} \circ u \circ g^{-1}: C \rightarrow A \Rightarrow f^{-1} \circ u \circ g^{-1} \in A^C$ and $h(f^{-1} \circ u \circ g^{-1}) = f \circ (f^{-1} \circ u \circ g^{-1}) \circ g = u$ proving surjectivity. $\square$

As we have that $\mathcal{P}(A) \approx 2^A$ and $\mathcal{P}(B) \approx 2^B$ and clearly $2 \approx 2$ we have the following corollary from the above theorem

Theorem 5.14. *If A, B are sets such that $A \approx B$ then $\mathcal{P}(A) \approx \mathcal{P}(B)$ and $2^A \approx 2^B$*

5.2 Infinite and finite sets

First every natural number is in reality a set this is expressed in the following definition and theorem.

Definition 5.15. *If $n \in \mathbb{N}_0$ then $S_n = \{m \in \mathbb{N}_0 | m < n\}$ (this is a initial segment of $\mathbb{N}_0$ (see 2.147))*

Theorem 5.16. *If $n \in \mathbb{N}_0$ then $n = S_n$*

Proof. Define $A = \{n \in \mathbb{N}_0 | n = S_n\}$ then we have

1. If $x \in S_0 \Rightarrow 0 \leqslant x < 0 \Rightarrow 0 < 0$ a contradiction so we have $S_0 = \emptyset = 0 \Rightarrow 0 \in A$
2. If $n \in A$ then $n = S_n$ if we take then $s(n) = s(S_n) = S_n \bigcup \{S_n\} = S_n \bigcup \{n\}$. Now if $m \in s(n) \Rightarrow m \in S_n$ or $m = n \Rightarrow m < n$ or $m = n < s(n) \Rightarrow m \in S_{s(n)}$. If $m \in S_{s(n)}$ then $m < s(n)$ we have then either
   a. $\boldsymbol{(m < n)}$ $\Rightarrow m \in S_n \Rightarrow m \in s(n)$
   b. $\boldsymbol{(n < m)}$ $\Rightarrow s(n) \leqslant m < s(n) \Rightarrow s(n < s(n))$ a contradiction, so this case can not occur.
   c. $\boldsymbol{(n = m)}$ $\Rightarrow m \in \{n\} \Rightarrow m \in S_n \bigcup \{n\} \Rightarrow m \in s(n)$

   So we have proved that $s(n) = S_{s(n)}$ and thus that $s(n) \in A$

By the principle of induction (see 4.10) we have that $A = \mathbb{N}_0$ and thus $\forall n \in \mathbb{N}_0$ that $n = S_n$. □

Theorem 5.17. (Recursion on $\boldsymbol{S_{n+1}}$ - Step Form) *Let A be a set, $a \in A$ and $g: S_n \times A \rightarrow A$ a function. Then there exists a unique function $f: S_{n+1} \rightarrow A$ satisfying*

$$\begin{aligned} f(0) &= a \\ \forall i \in S_n \text{ we have } f(i+1) &= g(i, f(i)) \end{aligned}$$

Proof. *Define $g': \mathbb{N}_0 \times A \rightarrow A$ by $(i, x) \rightarrow g'(i, x) = \left\{ \begin{smallmatrix} g(i,x) \text{ if } i < n \\ x \text{ if } n \leqslant i \end{smallmatrix} \right.$ then by 4.26 there exists a function $f': \mathbb{N}_0 \rightarrow A$ such that*

$$\begin{aligned} f'(0) &= a \\ \forall i \in \mathbb{N}_0 \text{ we have } f'(s(i)) = f'(i+1) &= g'(i, f'(i)) \end{aligned}$$

Define now $f: S_{n+1} \rightarrow A$ by $f = f'_{|S_{n+1}}$ then we have

1. $f(0) \underset{0 \in S_{n+1} (\text{as} 0 < n+1)}{=} f'(0) = a$
2. *$\forall i \in S_n$ we have $i + 1 \in S_{n+1}$ and thus $f(i+1) = f'(i+1) = g'(i, f'(i)) \underset{i<n}{=} g(i, f'(i)) \underset{i \in S_n \subseteq S_{n+1}}{=} g(i, f(i))$*

so we proved that f is the sought for function. Now to prove that it is unique let h: $S_{n+1} \rightarrow A$ be such that :

1. $h(0) = a$
2. $\forall i \in S_n$ *we have* $h(i+1) = g(i, h(i))$

Take now $B = \{i \in \mathbb{N}_0 | i \notin S_{n+1} \vee (i \in S_{n+1} \wedge h(i) = f(i))\}$ *then we have*

1. $h(0) = a = f(0)$ *and thus* $0 \in B$
2. *If* $i \in B$ *then we have the following cases*
   a. $(\boldsymbol{s(i) \notin S_{n+1}})$ $\Rightarrow s(i) \in B$
   b. $(\boldsymbol{s(i) \in S_{n+1}})$ *then* $i+1 < n+1 \Rightarrow i < n \Rightarrow i \in S_n$ *and* $h(s(i)) = h(i+1) = g(i, h(i)) \underset{i \in B \wedge i \in S_n \subseteq S_{n+1} \Rightarrow h(i) = f(i)}{=} g(i, f(i)) = f(i+1) = f(s(i))$ *so* $s(i) \in B$

Using mathematical induction (see 4.10) we have that $B = \mathbb{N}_0$. *Now if* $i \in S_{n+1} \Rightarrow i \in \mathbb{N}_0 = B \underset{i \in S_{n+1}}{\Rightarrow} h(i) = f(i)$ *and thus* $h = f$ □

We can create a second form of the above theorem as follows

Theorem 5.18. (Recursion on $\mathbb{N}_0$ - Step Form) *Let A be a set $a \in A$ and g: $\mathbb{N}_0 \times A \rightarrow A$ a function then there exists a function $f: \mathbb{N}_0 \rightarrow A$ such that*

$$\begin{aligned} f(0) &= a \\ \forall i \in \mathbb{N}_0 \text{ we have } f(i+1) &= g(i, f(i)) \end{aligned}$$

Proof. This is trivial by using 4.26 and the fact that $s(n) = n + 1$ □

Corollary 5.19. *Let A be a set, $n \in \mathbb{N}_0$, $a \in A$ and g: $\{n....\} \times A \rightarrow A$ a function then there exists a function $f: \{n, ...\} \rightarrow A$ such that*

$$\begin{aligned} f(n) &= a \\ \forall i \in \{n, ...\} \text{ we have } f(i+1) &= g(i, f(i)) \end{aligned}$$

Proof. Take g': $\mathbb{N}_0 \times A \rightarrow A$ by $(i, a) \rightarrow g(i+n, a)$ then by the previous theorem there exists a f': $\mathbb{N}_0 \rightarrow A$ such that

$$\begin{aligned} f'(0) &= a \\ \forall i \in \mathbb{N}_0 \text{ we have } f'(i+1) &= g'(i, f'(i)) \end{aligned}$$

Take now f: $\{n, ...\} \rightarrow A$ by $i \rightarrow f'(i-n)$ then we have

$$\begin{aligned} f(n) &= f'(n-n) = f'(0) = a \\ \forall i \in \{n, ...\} \text{ we have } i - n \in \mathbb{N}_0 \text{ and } f(i+1) &= f'((i-n)+1) = g'((i-n), f'(i-n)) = g(i, f(i)) \end{aligned}$$

□

Example 5.20. Define $g: \mathbb{N}_0 \times \mathbb{N}_0 \rightarrow \mathbb{N}_0$ by $g(i, n) = (i+1) \cdot n$ and take $a = 1$ then using the above theorem there exists a function $f: \mathbb{N}_0 \rightarrow \mathbb{N}_0$ such that

$$\begin{aligned} f(0) &= 1 \\ \forall i \in \mathbb{N}_0 \text{ we have } f(i+1) &= (i+1) \cdot f(i) \end{aligned}$$

We note $f(i)$ as $i!$ called the faculty and as some values we have

$$\begin{aligned} 0! &= 1 \\ 1! &= (1+0)! \\ & \quad (1+0) \cdot 0! = 1 \\ 2! &= (1+1)! \\ &= 2 \cdot 1! = 2 \\ 3! &= 3 \cdot 2! = 6 \\ & \ldots \end{aligned}$$

The step form of recursion leads to the following corollary using finite sequences and sequences.

Corollary 5.21. *Let A be a set, $a \in A$, $\mathcal{M} \subseteq \{\{x_i\}_{i \in \{0, \ldots, n\}} | n \in \mathbb{N}_0 \wedge x_i \in A\}$ [the set of finite families of elements in A (or the set of function from $S_n \rightarrow A$, $n \in \mathbb{N}$] and $\rho: \mathcal{M} \rightarrow A$ a function then there exists a function $f: \mathbb{N}_0 \rightarrow A$ such that*

1. $f(0) = a$
2. $\forall n \in \mathbb{N}_0$ *we have* $f(s(n)) = \rho(\{f(i)\}_{i \in \{0, \ldots, n\}})$

Proof. Let $\{c_i\}_{i \in \{0, \ldots, 0\}}$ be defined by $c_0 = a$ and define now $g: \mathcal{M} \rightarrow \mathcal{M}$ by $\{x_i\}_{i \in \{0, \ldots, n\}} \rightarrow g(\{x_i\}_{i \in \{0, \ldots, n\}})$ where $\forall i \in \{0, \ldots, n+1\}$ we have $g(\{x_i\}_{i \in \{0, \ldots, n\}}) = \begin{cases} x_i \text{ if } 0 \leqslant i \leqslant n \\ \rho(\{x_i\}_{i \in \{0, \ldots, n\}}) \text{ if } i = n+1 \end{cases}$ using recursion (see 4.18) we have then the existence of a function $h: \mathbb{N}_0 \rightarrow \mathcal{M}$ where

1. $h(0) = \{c_i\}_{i \in \{0, \ldots, 0\}}$
2. $\forall n \in \mathbb{N}_0$ we have $h(n+1) = g(h(n))$

Define then $S = \{n \in \mathbb{N}_0 | \mathrm{dom}(h(n)) = \{0, \ldots, n\}\}$ then we have

1. If $n = 0$ then $\mathrm{dom}(h(0)) = \mathrm{dom}(\{c_i\}_{i \in \{0, \ldots, 0\}}) = \{0, \ldots 0\}$
2. If $n \in S$ then $\mathrm{dom}(h(n+1)) = \mathrm{dom}(g(h(n))) \underset{\text{definition of } h}{=} \{0, \ldots, \mathrm{dom}(h(n)) + 1\} \underset{n \in S}{=} \{0, \ldots n+1\}$

Using mathematical induction we have thus that $S = \mathbb{N}_0 \Rightarrow \forall n \in \mathbb{N}_0$ that $\mathrm{dom}(h(n)) = \{0, \ldots, n\}$. Define now $F = \{n \in \mathbb{N}_0 | \text{if } i \leqslant n \wedge j \leqslant i \text{ then } h(n)_i = h(i)_j\}$ then we have

1. If $n = 0$ then from $i \leqslant n \wedge j \leqslant i$ we have $i = j = 0$ then $h(n)_j = h(0)_0 = h(i)_j$ or $0 \in F$
2. If $n \in F$ then we have that $\mathrm{dom}(h(n+1)) = \{0, \ldots, n+1\}$ and if $i \leqslant n+1 \wedge j \leqslant i$ we have the following cases
   a. $\boldsymbol{(i = n+1)}$ then $h(n+1)_j = h(i)_j$

b. **($i < n + 1$)** then $i \leqslant n$ and if $j \leqslant i$ we have $h(n+1)_j = g(h(n))_j \underset{j<n+1+\text{definition of } g}{=} h(n)_j \underset{n \in F}{=} h(i)_j$

So in both cases we have $h(n+1)_j = h(i)_j$ proving that $n+1 \in F$

Using mathematical induction we have then that $F = \mathbb{N}_0$ or $\forall n \in \mathbb{N}_0$ we have then if $i \leqslant n$ and $j \leqslant i$ then $h(n)_j = h(i)_j$. Define now $f: \mathbb{N}_0 \to A$ by $f(n) = h(n)_n$ then we have

1. $f(0) = h(0)_0 = c_0 = a$
2. $f(n+1) = h(n+1)_{n+1} = g(h(n))_{n+1} = \rho(h(n))$, where if $k \in \text{dom}(h(n)) = \{0, \ldots, n\}$ we have that $h(n)_k \underset{k \leqslant n, k \leqslant k}{=} h(k)_k = f(k)$ so that $h(n) = \{f(i)\}_{i \in \{0, \ldots, n\}}$ giving that $f(n+1) = \rho(\{f_i\}_{i \in \{1, \ldots, n\}})$ □

Lemma 5.22. *Let A be a set $\mathcal{M} = \{\{x_i\}_{i \in \{0, \ldots, n\}} | n \in \mathbb{N}_0 \wedge \forall i \in \{0, \ldots, n\}$ we have $x_i \in A\}$, $m \in \mathbb{N}_0$ and $\mathcal{M}_m = \{\{x_i\}_{i \in \{m, \ldots, n\}} | n \in \{m, \ldots\} \wedge \forall i \in \{m, \ldots, n\}$ we have $x_i \in A\}$. Define then $T_m: \mathcal{M} \to \mathcal{M}_m$ by $\{x_i\}_{i \in \{0, \ldots, n\}} \to T_m(\{x_i\}_{i \in \{0, \ldots, m\}}) = \{(T_m(\{x_i\}_{i \in \{0, \ldots, n\}}))_j\}_{j \in \{m, \ldots, n+m\}}$ where for $j \in \{m, \ldots, n+m\}$ we have $(T_m(\{x_i\}_{i \in \{0, \ldots, \}}))_j = x_{j-m}$ then we have*

1. *T_m is a bijection*
2. *The inverse of T_m is $T_{-m}: \mathcal{M}_m \to \mathcal{M}$ defined by $\{x_i\}_{i \in \{m, \ldots, n\}} \to \{(T_{-m}(\{x_i\}_{i \in \{m, \ldots, n\}}))_j\}_{j \in \{0, \ldots, n-m\}}$ where if $j \in \{0, \ldots, n-m\}$ then $(T_{-m}(\{x_i\}_{i \in \{m, \ldots, n\}}))_j = x_{j+m}$*

As a shorthand we note $\{T_m(\{x_i\}_{i \in \{0, \ldots, n\}})\}_{j \in \{m, \ldots, n+m\}} \underset{\text{notation}}{=} \{x_{i-m}\}_{i \in \{m, \ldots, n+m\}}$ and $\{T_m(\{x_i\}_{i \in \{m, \ldots, n\}})\}_{j \in \{0, \ldots, n-m\}} \underset{\text{notation}}{=} \{x_{i+m}\}_{i \in \{0, \ldots, n-m\}}$.

Proof.

1. **(injectivity)** If $\{x_i\}_{i \in \{0, \ldots, n_1\}}, \{y_i\}_{i \in \{0, \ldots, n_2\}} \in \mathcal{M}$ then if $T_m(\{x_i\}_{i \in \{0, \ldots, n_1\}}) = T_m(\{y_i\}_{i \in \{0, \ldots, n_2\}})$ we must have that $n_1 + m = n_2 + m \Rightarrow n_1 = n_2 = n$ (consider families as functions which must be equal so their domain must be equal) and $\forall i \in \{0, \ldots, n\}$ we have $i + m \in \{m, \ldots, n+m\}$ so that $x_i = x_{(i+m)-m} = (T_m(\{x_i\}_{i \in \{0, \ldots, n\}}))_{i+m} = (T_m(\{y_i\}_{i \in \{0, \ldots, n\}}))_{i+m} = y_{(i+m)-m} = y_i$ proving that $\{x_i\}_{i \in \{0, \ldots, n_1\}} = \{y_i\}_{i \in \{0, \ldots, n_2\}}$
2. **(surjectivity)** Let $\{x_i\}_{i \in \{m, \ldots, n\}} \in \mathcal{M}_m$ define then $\{y_i\}_{i \in \{0, \ldots, n-m\}}$ by $\forall i \in \{0, \ldots, n-m\}$ we have $y_i = x_{i+m}$ then $\forall i \in \{m, \ldots, n\}$ we have $(T_m(\{y_i\}_{i \in 0, \ldots, n-m}))_i = y_{i-m} = x_{(i+m)-m} = x_i$ proving that $T_m(\{y_i\}_{i \in 0, \ldots, n-m}) = \{x_i\}_{i \in m, \ldots, n}$

Finally to prove that T_{-m} is the inverse of T_m, take $\{x_i\}_{i \in \{0, \ldots., n\}}$ then $T_{-m}(T_m(\{x_i\}_{i \in \{0, \ldots, n\}})) = T_{-m}(\{(T_m(x_i)_{i \in \{0, \ldots, n\}})_j\}_{j \in \{m, \ldots, m+n\}}) = \{(T_{-m}(\{(T_m(x_i)_{i \in \{0, \ldots, n\}})_j\}_{j \in \{m, \ldots, m+n\}}))_k\}_{k \in \{0, \ldots, (m+n)-m\}} = \{(T_{-m}(\{x_{j-m}\}_{j \in \{m, \ldots, m+n\}}))_k\}_{k \in \{0, \ldots, n\}} = \{x_{(k-m)+m}\}_{k \in \{0, \ldots, n\}} = \{x_k\}_{k \in \{0, \ldots, k\}}$ □

Corollary 5.23. *Let A be a set, $a\in A$,$m\in\mathbb{N}_0$, $\mathcal{M}_m=\{\{x_i\}_{i\in\{m,\ldots,n\}}|n\in\{m,\ldots,\}\wedge x_i\in A\}$ [the set of finite families of elements in A and $\rho{:}\,\mathcal{M}_m\rightarrow A$ a function then there exists a function $f{:}\,\{m,\ldots\}\rightarrow A$ such that*

1. *$f(m)=a$*
2. *$\forall n\in\{m,\ldots\}$ we have $f(n+1)=f(s(n))=\rho(\{f(i)\}_{i\in\{m,\ldots,n\}})$*

Proof. If we define $\mathcal{M}=\{\{x_i\}_{i\in\{0,\ldots,n\}}|n\in\mathbb{N}_0\wedge\forall i\in\{0,\ldots,n\}\text{ we have }x_i\in A\}$ then using the previous lemma (see 5.22) we have the bijections $T_m{:}\,\mathcal{M}\rightarrow\mathcal{M}_m$ and $T_{-m}{:}\,\mathcal{M}_m\rightarrow\mathcal{M}$, define now $\rho'{:}\,\mathcal{M}\rightarrow A$ by $\rho'=\rho\circ T_m$ then $\rho'{:}\,\mathcal{M}\rightarrow A$ and by 5.21 there exists a $f'{:}\,\mathbb{N}_0\rightarrow A$ such that

1. $f'(0)=a$
2. $f'(n+1)=\rho'(\{f(i)\}_{i\in\{0,\ldots,n\}})$

Define now $f{:}\,\{m,\ldots\}\rightarrow A$ by $f(i)=f'(i-m)$ so that if $j\in\mathbb{N}_0$ then $f'(j)=f(j+m)$ then we have

1. $f(m)=f'(m-m)=f'(0)=a$
2. $\forall n\in\{m,\ldots\}$ we have $f(n+1)=f'((n+1)-m)=f'((n-m)+1)=\rho'(\{f'(i)_{i\in\{0,\ldots,n-m\}}\})=\rho(T_m(\{f'(i)\}_{i\in\{0,\ldots,n-m\}}))=\rho(T_m(\{f(i+m)\}_{i\in\{0,\ldots,n-m\}}))=\rho(T_m(T_{-m}(\{f(i)\}_{i\in\{m,\ldots,n\}})))=\rho(\{f(i)\}_{i\in\{m,\ldots,n\}})$ □

Corollary 5.24. *Let A be a set, $a\in A$, $m\in\mathbb{N}_0$, $\mathcal{M}=\{\{x_i\}_{i\in\{m,\ldots,n\}}|n\in\{m,\ldots\}\wedge\forall i\in\{m,\ldots,n\}\text{ we have }x_i\in A\}$ [the set of finite families of elements in A] and $\rho{:}\,\mathcal{M}\rightarrow A$ a function so that $\forall\{x_i\}_{i\in\{m,\ldots,n\}}\in\mathcal{M}$ we have that $P(\{x_i\}_{i\in\{m,\ldots,n\}},\rho(\{x_i\}_{i\in\{m,\ldots,n\}}))$ is true then there exists a function $f{:}\,\{m,\ldots\}\rightarrow A$ such that*

1. *$f(m)=a$*
2. *$\forall n\in\mathbb{N}_0$ we have $f(s(n))=\rho(\{f(i)\}_{i\in\{m,\ldots,n\}})$*
3. *$\forall n\in\mathbb{N}_0$ we have that $P(\{f(i)\}_{i\in\{m,\ldots,n\}},f(n+1))$ is true*

Proof. First we use the previous corollary to prove that there exists a $f{:}\,\{m,\ldots\}\rightarrow A$ such that

1. $f(m)=a$
2. $\forall n\in\{m,\ldots\}$ we have $f(s(n))=\rho(\{f(i)\}_{i\in\{m,\ldots,n\}})$

We proceed now by induction so let $S=\{n\in\{m,\ldots\}|P(\{f(i)\}_{i\in\{m,\ldots,n\}},f(n+1))\text{ is true}\}$ then we have

1. if $n=m$ then as $P(\{f(i)\}_{i\in\{m,\ldots,m\}},f(m+1))=P(\{f(i)_{i\in\{m,\ldots,m\}}\},\rho(\{f(i)\}_{i\in\{m,\ldots,m\}}))$ is true by the hypothesis so that $m\in S$
2. If $n\in S$ then by assumption we have that $P(\{f(i)\}_{i\in\{m,\ldots,n\}},\rho(\{f(i)\}_{i\in\{m,\ldots n\}}))=P(\{f(i)\}_{i\in\{m,\ldots,n\}},f(n+1))$ so we have that $n+1\in S$

Which by mathematical induction proves the theorem. □

Another variant of recursion and families is the following

Theorem 5.25. *Let A be a set and $\mathcal{M}=\{\{x_i\}_{i\in\{0,\ldots,n\}}|\forall i\in\{0,\ldots,n\}$ we have $x_i\in A\}$, $\mathcal{N}\subseteq\mathcal{M}$ and $\rho:\mathcal{N}\rightarrow A$ such that for $\rho(\{x_i\}_{i\in\{0,\ldots,n\}})$ we have $\rho(\{x_i\}_{i\in 0,\ldots,n})\in A$ and $\{x'_i\}_{i\in\{0,\ldots,n+1\}}\in\mathcal{N}$ (where $x'_i=\begin{cases} x_i \text{ is } i\in\{0,\ldots,n\} \\ \rho(\{x_i\}_{i\in\{0,\ldots,n\}}) \text{ if } i=n+1\end{cases}$. Then for $a\in A$ with $\{a\}_{i\in\{0,\ldots,0\}}\in\mathcal{N}$ there exists a function $f:\mathbb{N}_0\rightarrow A$ such that*

1. $f(0)=a$
2. $\forall n\in\mathbb{N}_0$ *we have* $f(n+1)=\rho(\{f(i)\}_{i\in\{0,\ldots,n\}})$
3. $\forall n\in\mathbb{N}_0$ *we have* $\{f(i)\}_{i\in\{0,\ldots,n\}}\in\mathcal{N}$

Proof. Take $\{a\}_{i\in\{0,\ldots,0\}}\in\mathcal{N}$ (by the hypothesis) and define $g:\mathcal{N}\rightarrow\mathcal{N}$ by $\{x_i\}_{i\in\{0,\ldots,n\}}\rightarrow g(\{x_i\}_{i\in\{0,\ldots,n\}})$ where $\forall i\in\{0,\ldots,n+1\}$ we have $g(\{x_i\}_{i\in\{0,\ldots,n\}})=\begin{cases} x_i \text{ if } 0\leqslant i\leqslant n \\ \rho(\{x_i\}_{i\in\{0,\ldots,n\}}) \text{ if } i=n+1\end{cases}\in\mathcal{N}$(by the hypothesis), using recursion (see 4.18) we have then the existence of a function $h:\mathbb{N}_0\rightarrow\mathcal{N}$ where

1. $h(0)=\{a\}_{i\in\{0,\ldots,0\}}$
2. $\forall n\in\mathbb{N}_0$ we have $h(n+1)=g(h(n))$

Define then $S=\{n\in\mathbb{N}_0|\mathrm{dom}(h(n))=\{0,\ldots,n\}\}$ then we have

1. If $n=0$ then $\mathrm{dom}(h(0))=\mathrm{dom}(\{a\}_{i\in\{0,\ldots,0\}})=\{0,\ldots 0\}$
2. If $n\in S$ then $\mathrm{dom}(h(n+1))=\mathrm{dom}(g(h(n)))\underset{\text{definition of } h}{=}\{0,\ldots,\mathrm{dom}(h(n))+1\}\underset{n\in S}{=}\{0,\ldots n+1\}$

Using mathematical induction we have thus that $S=\mathbb{N}_0\Rightarrow\forall n\in\mathbb{N}_0$ that $\mathrm{dom}(h(n))=\{0,\ldots,n\}$. Define now $F=\{n\in\mathbb{N}_0|\text{if } i\leqslant n\wedge j\leqslant i \text{ then } h(n)_i=h(i)_j\}$ then we have

1. If $n=0$ then from $i\leqslant n\wedge j\leqslant i$ we have $i=j=0$ then $h(n)_j=h(0)_0=h(i)_j$ or $0\in F$
2. If $n\in F$ then we have that $\mathrm{dom}(h(n+1))=\{0,\ldots,n+1\}$ and if $i\leqslant n+1\wedge j\leqslant i$ we have the following cases
   a. $(\boldsymbol{i=n+1})$ then $h(n+1)_j=h(i)_j$
   b. $(\boldsymbol{i<n+1})$ then $i\leqslant n$ and if $j\leqslant i$ we have $h(n+1)_j=g(h(n))_j\underset{j<n+1+\text{definition of } g}{=}h(n)_j\underset{n\in F}{=}h(i)_j$

   So in both cases we have $h(n+1)_j=h(i)_j$ proving that $n+1\in F$

Using mathematical induction we have then that $F=\mathbb{N}_0$ or $\forall n\in\mathbb{N}_0$ we have then if $i\leqslant n$ and $j\leqslant i$ then $h(n)_j=h(i)_j$. Define now $f:\mathbb{N}_0\rightarrow A$ by $f(n)=h(n)_n$ then we have

1. $f(0)=h(0)_0=c_0=a$
2. $f(n+1)=h(n+1)_{n+1}=g(h(n))_{n+1}=\rho(h(n))$, where if $k\in\mathrm{dom}(h(n))=\{0,\ldots,n\}$ we have that $h(n)_k\underset{k\leqslant n,k\leqslant k}{=}h(k)_k=f(k)$ so that $h(n)=\{f(i)\}_{i\in\{0,\ldots,n\}}$ giving that $f(n+1)=\rho(\{f_i\}_{i\in\{1,\ldots,n\}})$

Let $\mathcal{W}=\{n\in\mathbb{N}_0|\{f(i)_{i\in\{0,\ldots,n\}}\}\in\mathcal{N}\}$ then we have:

1. $\{f(i)_{i\in\{0,\ldots,0\}}\}=\{f(0)\}_{i\in\{0,\ldots,0\}}=\{a\}_{i\in\{0,\ldots,0\}}\in\mathcal{N}$

2. If $n \in \mathcal{W}$ then $\{f(i)\}_{i \in \{0, \ldots, n+1\}} = \rho(\{f(i)_{i \in \{0, \ldots, n+1\}}\}) \in \mathcal{N}$ as $\{f(i)\}_{i \in \{0, \ldots, n\}} \in \mathcal{N}\, [n \in \mathcal{W}]$ and the hypothesis about ρ.

Applying mathematical induction on $\mathcal{W}$ gives then $\mathcal{W} = \mathbb{N}_0$ and thus (3) in our theorem. □

Corollary 5.26. *Let A be a set and $\mathcal{M}_m = \{\{x_i\}_{i \in \{m, \ldots, n\}} | n \in \{m, \ldots\} \wedge \forall i \in \{0, \ldots, n\}$ we have $x_i \in A\}$, $\mathcal{N}_m \subseteq \mathcal{M}_m$ and $\rho: \mathcal{N}_m \rightarrow A$ such that for $\rho(\{x_i\}_{i \in \{m, \ldots, n\}})$ we have $\rho(\{x_i\}_{i \in m, \ldots, n}) \in A$ and $\{x_i'\}_{i \in \{m, \ldots, n+1\}} \in \mathcal{N}_m$ (where $x_i' = \begin{cases} x_i \text{ is } i \in \{m, \ldots, n\} \\ \rho(\{x_i\}_{i \in \{m, \ldots, n\}}) \text{ if } i = n+1 \end{cases}$. Then for $a \in A$ with $\{a\}_{i \in \{m, \ldots, m\}} \in \mathcal{N}_m$ there exists a function $f: \mathbb{N}_0 \rightarrow A$ such that*

1. *$f(m) = a$*
2. *$\forall n \in \{m, \ldots\}$ we have $f(n+1) = \rho(\{f(i)\}_{i \in \{m, \ldots, n\}})$*
3. *$\forall n \in \{m, \ldots\}$ we have $\{f(i)\}_{i \in \{m, \ldots, n\}} \in \mathcal{N}$*

Proof. If we define $\mathcal{M} = \{\{x_i\}_{i \in \{0, \ldots, n\}} | n \in \mathbb{N}_0 \wedge \forall i \in \{0, \ldots, n\}$ we have $x_i \in A\}$ then using 5.22 we have the bijections $T_m: \mathcal{M} \rightarrow \mathcal{M}_m$ and $T_{-m}: \mathcal{M}_m \rightarrow \mathcal{M}$. Define then $\mathcal{N} = T_{-m}(\mathcal{N}_m)$ so that we have the bijections $T_{-m|\mathcal{N}_m}: \mathcal{N}_m \rightarrow \mathcal{N}$ and $T_{m|\mathcal{N}}: \mathcal{N} \rightarrow \mathcal{N}_m$. Define then $\rho': \mathcal{N} \rightarrow A$ as $\rho' = \rho \circ T_{m|\mathcal{N}}$ then if $\{x_i\}_{i \in \{0, \ldots, n\}} \in \mathcal{N}$ we have $\rho'(\{x_i\}_{i \in \{0, \ldots, n\}}) = \rho(T_m(\{x_i\}_{i \in \{0, \ldots, n\}})) = \rho(\{x_{i-m}\}_{i \in \{m, \ldots, n+m\}})$. If we then create $\{x_i'\}_{i \in \{0, \ldots, n+1\}} \in \mathcal{N}$ by $x_i' = \begin{cases} x_i \text{ if } i \in \{0, \ldots, n\} \\ \rho'(\{x_i\}_{i=n+1}) \end{cases} = \begin{cases} x_i \text{ if } i \in \{0, \ldots, n\} \\ \rho(T_m(\{x_i\}_{i \in \{0, \ldots, n\}})) \end{cases}$ then we have for $T_{m|\mathcal{N}}(\{x_i'\}_{i \in \{0, \ldots, n+1\}})$ that if $k \in \{m, \ldots, n+m+1\}$ $(T_{m|\mathcal{N}_m}(\{x_i'\}_{i \in \{0, \ldots, n+1\}}))_k = x_{j-m}' = \begin{cases} T_m(\{x_i\}_{i \in \{0, \ldots, n\}})_k \text{ if } k \in \{m, \ldots, n+m\} \\ \rho(T_m(\{x_i\}_{i \in \{0, \ldots, n\}})) \text{ if } k = n+m+1 \end{cases}$ so by the properties of ρ we have that $T_m(\{x_i'\}_{i \in \{0, \ldots, n+1\}}) \in \mathcal{N}_m$ so that $\{x_n'\}_{i \in \{0, \ldots, n+1\}} \in \mathcal{N}$. Also using the fact that $\{a\}_{i \in \{m, \ldots, m\}} \in \mathcal{N}_m$ we have that $\{a\}_{i \in \{0, \ldots, 0\}} = T_{-m}(\{a\}_{i \in \{0, \ldots, 0\}}) \in \mathcal{N}$ Using the above theorem (see 5.25) there exists then a $f': \mathbb{N}_0 \rightarrow A$ such that

1. $f'(0) = a$
2. $\forall n \in \mathbb{N}_0$ we have $f'(n+1) = \rho'(\{f'(i)\}_{i \in \{0, \ldots, n\}})$
3. $\forall n \in \mathbb{N}_0$ we have $\{f'(i)\}_{i \in \{0, \ldots, n\}} \in \mathcal{N}$

Define then $f: \{m, \ldots\} \rightarrow A$ by $i \rightarrow f(i) = f'(i-m)$ or if $i \in \mathbb{N}_0$ then $f'(i) = f(i+m)$ then we have

1. $f(m) = f'(m-m) = f'(0) = a$
2. $\forall n \in \{m, \ldots\}$ we have $f(n+1) = f'((n-m)+1) = \rho'(\{f'(i)\}_{i \in \{0, \ldots, (n-m)\}}) = \rho'(\{f(i+m)\}_{i \in \{0, \ldots, (n-m)\}}) = \rho'(T_{-m}(\{f_i\}_{i \in \{m, \ldots, n\}})) = \rho(T_m(T_{-m}(\{f_i\}_{i \in \{m, \ldots, n\}}))) = \rho(\{f(i)\}_{i \in \{m, \ldots, n\}})$
3. $\forall n \in \{m, \ldots\}$ we have $\{f'(i)\}_{i \in \{0, \ldots n-m\}} \in \mathcal{N}$ so that $T_m(\{f'(i)\}_{i \in \{0, \ldots, n-m\}}) \in T_m(T_{-m}(\mathcal{N}_m)) = \mathcal{N}_m$, as $T_m(\{f'(i)\}_{i \in \{0, \ldots, n-m\}}) = \{f'(i-m)\}_{i \in \{m, \ldots, n\}} = \{f(i)\}_{i \in \{m, \ldots, n\}}$ this means that $\{f(i)\}_{i \in \{m, \ldots, n\}} \in \mathcal{N}$ □

Definition 5.27. *A set A is called finite if $\exists n \in \mathbb{N}_0$ such that $n \approx A$. A set that is not finite is called a infinite set.*

Note 5.28. $\emptyset: \emptyset \rightarrow 0 = \emptyset$ is a bijection as all the conditions of a bijection are vacuously true, so the empty set is finite.

Notation 5.29. *If $A \neq \emptyset$ is a finite set then there exists a bijection $a: S_{n+1} \rightarrow A$ and we note then $A = \{a_0, \ldots, a_n\}$ where $a_i = a(i)$ (note that sometimes if $n \in \mathbb{N}$ we write $A = \{b_1, \ldots, b_n\}$ where $b_i = a(i+1)$ and $a: S_n \rightarrow A$ is a bijection)*

Definition 5.30. *A set A is called **denumerable** or **infinite countable** if $A \approx \mathbb{N}_0$*

Definition 5.31. *A set A is called **countable** if it is either **finite** or **infinite countable***

Lemma 5.32. *If A is a **denumerable** set and $x \in A$ then $A \backslash \{x\}$ is a **denumerable** set.*

Proof. If A is countable then $A \approx \mathbb{N}_0$ and thus there exists a bijection $f: \mathbb{N}_0 \rightarrow A$. By surjectivity of f and $x \in A$ there exists a $n \in \mathbb{N}_0$ such that $f(n) = x$. Define now $g: \mathbb{N}_0 \rightarrow A \setminus \{x\}$ by $g(m) = \begin{cases} f(m) \text{ if } m < n \\ f(m+1) \text{ if } n \leqslant m \end{cases}$. We prove now that g is bijective:

1. **(injectivity)** If $g(m) = g(m')$ then we have the following possible cases
   a. $(\boldsymbol{m < n \wedge m' < n})$ then we have $f(m) = g(m) = g(m') = f(m') \Rightarrow f(m) = f(m') \underset{f \text{ is injective}}{\Rightarrow} m = m'$
   b. $(\boldsymbol{m < n \wedge n \leqslant m'})$ then we have $f(m) = g(m) = g(m') = f(m'+1) \Rightarrow f(m) = f(m'+1) \Rightarrow m = m'+1$. Then as $n \leqslant m' \Rightarrow n < n+1 \leqslant m'+1 = m$ we have from $m < n$ that $m < m$ a contradiction, so this case does not apply.
   c. $(\boldsymbol{n \leqslant m \wedge m' < n})$ then we have $f(m+1) = g(m) = g(m') = f(m') \Rightarrow f(m+1) = f(m') \Rightarrow m+1 = m' \underset{n \leqslant m}{\Rightarrow} n+1 \leqslant m+1 = m' \underset{m' < n}{\Rightarrow} n+1 < n < n+1 \Rightarrow n+1 < n+1$ again a contradiction, so we can ignore this case as it never occurs.
   d. $(\boldsymbol{n \leqslant m \wedge n \leqslant m'})$ then $f(m+1) = g(m) = g(m') = f(m'+1) \Rightarrow f(m+1) = f(m'+1) \Rightarrow m+1 = m'+1 \Rightarrow m = m'$
2. **(surjectivity)** If $y \in A \backslash \{x\}$ then there exists by surjectivity of g a $i \in \mathbb{N}_0$ such that $f(i) = y$. We can not have that $i = n$ as then $y = f(i) = x \notin A \backslash \{x\}$. So the only cases for i are
   a. $(\boldsymbol{i < n})$ then $g(i) = f(i) = y \Rightarrow g(i) = y$
   b. $(\boldsymbol{n < i})$ then $n+1 \leqslant i \Rightarrow n \leqslant i-1 \Rightarrow g(i-1) = f((i-1)+1) = f(i) = y \Rightarrow g(i-1) = y$

   proving surjectivity. □

Lemma 5.33. *If $n \in \mathbb{N}_0$ then n has no **denumerable** subset*

Proof. We prove this by induction on n, so let $S = \{n \in \mathbb{N}_0 | n \text{ does not contain } a \text{ enumerable subset}\}$.

1. First as $0 = \emptyset$ we have if $A \subseteq 0 = \emptyset \Rightarrow \emptyset \subseteq A \subseteq \emptyset \Rightarrow A = \emptyset$. If now $\mathbb{N}_0 \approx A$ then there exists a bijection $f: \mathbb{N}_0 \rightarrow A$ and thus $f(0) \in A \Rightarrow A \neq \emptyset$ a contradiction. So 0 does not contains a denumerable subset and we have $0 \in S$.

2. Assume now that $n \in S$. We proceed now by contradiction, so assume that there exists a $A \subseteq n+1 = s(n) = n \bigcup \{n\}$ such that A is denumerable. If $n \notin A \Rightarrow A \subseteq n$ which is impossible as $n \in S \Rightarrow n \in A$. Now if $x \in A \backslash \{n\} \subseteq A \subseteq n \bigcup \{n\}$ then $x \in n \bigcup \{n\}$ and from $x \neq n$ we have that $x \in n$ so we have $A \backslash \{n\} \subseteq n$ and as by the previous lemma we would have that $A \backslash \{n\}$ is denumerable, we reach the contradiction that $n \notin S$. So we can not have a countable subset of $n+1$ and thus $n+1 \in S$

Using 4.10 we conclude that $S = \mathbb{N}_0$ proving our theorem. □

Theorem 5.34. *If A is a infinite set if and only A has a denumerable subset. So any denumerable set must be infinite.*

Proof. First we prove that if A is infinite then A has a denumerable subset. Using the 'well ordering theorem' which is derived from the axiom of choice (see 2.218) there exists a order $\leqslant_A$ on A such that $\langle A, \leqslant_A \rangle$ is a well ordered set. Using 2.195 and the fact that $\langle \mathbb{N}_0, \leqslant \rangle$ is also well ordered (see 4.52) we have exactly one of the following:

1. **($\langle \mathbb{N}_0, \leqslant \rangle$ is isomorphic with $\langle A, \leqslant_A \rangle$)** but this would mean that $A \approx \mathbb{N}_0$ and thus A has a denumerable subset.

2. **($\langle \mathbb{N}_0, \leqslant \rangle$ is isomorphic with an initial segment of $\langle A, \leqslant_A \rangle$)** but this would again mean that A has a denumerable subset.

3. **($\langle A, \leqslant_A \rangle$ is isomorphic with an initial segment of $\langle \mathbb{N}_0, \leqslant \rangle$)** But then there exists a $n \in \mathbb{N}_0$ so that $A \approx S_n = n$ and thus A is finite which contradicts the fact that A is infinite, so this case does not apply.

Second assume that A has a denumerable subset B. If A is finite then there exists a $n \in \mathbb{N}_0$ such that $A \approx n$ and thus there exists a bijection $f: A \rightarrow n$. Then $f(B) \subseteq n$ and as $f_{|B}: B \rightarrow f(B)$ is a bijection and as $\mathbb{N}_0 \approx B$ we have the existence of a bijection $g: \mathbb{N}_0 \rightarrow B$ so $f_{|B} \circ g: \mathbb{N}_0 \rightarrow f(B)$ is a bijection and thus $f(B)$ is a denumerable subset of n which is impossible by the previous lemma. we must thus conclude that A is not finite and thus is infinite. □

Corollary 5.35. *Every set with a infinite subset is infinite.*

Proof. If A is a set such that there exists a infinite set B with $B \subseteq A$ then as B is infinite we have by 5.34 the existence of a denumerable set $C \subseteq B$ but then $C \subseteq A$ and thus A has a denumerable subset and by 5.34 we have that A is infinite. □

Corollary 5.36. *Every subset of a finite set is finite.*

Proof. If a finite set would contain a infinite subset then by the previous theorem the finite set would be infinite. □

Theorem 5.37. *If A, B are finite sets then $A \bigcup B$ is finite*

Proof. As A is finite we have by 5.36 that $A \backslash B$ is finite, this together with the finiteness of B give the existence of two bijections $f: A \backslash B \rightarrow n = S_n$ and g': $B \rightarrow m = S_m$ where $n, m \in \mathbb{N}_0$. Define now $g: B \rightarrow \{i \in \mathbb{N}_0 | n \leqslant i \wedge i < n+m\}$ by $g(x) = g'(x)+n$ then this function is bijective

1. **(injectivity)** If $g(x) = g(x')$ then $g'(x)+n = g'(x')+n \underset{4.44}{\Rightarrow} g'(x) = g'(x') \Rightarrow x = x'$
2. **(surjectivity)** If $i \in \{i \in \mathbb{N}_0 | n \leqslant i \wedge i < n+m\}$ then $n \leqslant i < n+m$ and thus using 4.62 there exists a $k \in \mathbb{N}_0$ such that $n+k = i$. We must have $k < m$ [if $m \leqslant k \Rightarrow n+m \leqslant n+k = i < n+m \Rightarrow$ we get the contradiction $n+m \neq n+m$] and thus $k \in S_m = m$. So using the fact that g' is bijective and thus surjective we have $\exists x \in B$ such that $g'(x) = k \Rightarrow g(x) = g'(x)+n = k+n = i$.

Further if $i \in n \bigcap \{i \in \mathbb{N}_0 | n \leqslant i \wedge i < n+m|\} \underset{n=S_n}{\Rightarrow} i < n \wedge n \leqslant i \Rightarrow i \neq i$ a contradiction, so we have $n \bigcap \{i \in \mathbb{N}_0 | n \leqslant i \wedge i < n+m\} = \emptyset$. So we can use 2.43 to construct the bijection $f \bigcup g: A \bigcup B = (A \backslash B) \bigcup B \rightarrow n \bigcup \{i \in \mathbb{N}_0 | n \leqslant i \wedge i < n+m|\} = S_{n+m} = n+m$ and thus we have proved that $A \bigcup B \approx n+m$. □

Lemma 5.38. *If $\{A_i\}_{i \in S_n}$ is a finite family of finite sets then $\bigcup_{i \in S_n} A_i$ is finite.*

Proof. We prove this by induction on n. So define $S = \{n \in \mathbb{N}_0 | \text{if } \{A_i\}_{i \in S_n} \text{ is } a \text{ family of finite set then } \bigcup_{i \in S_n} A_i \text{ is finite}\}$ then we have

1. If $n = 0$ take then $\{A_i\}_{i \in S_0}$ then there exists a graph A with $\operatorname{dom}(A) = \emptyset$ so if $x \in \bigcup_{i \in S_0} A_i$ then there exists a $i \in S_0$ such that $x \in A_i$ which is a contradiction as $S_0 = \emptyset$ so we have that $\bigcup_{i \in S_0} A_i = \emptyset$ and thus finite. So we have $0 \in S$
2. Assume that $n \in S$ and take $n+1$ then we have for the family of finite sets $\{A_i\}_{i \in S_{n+1}}$ that there exists a graph A with $\operatorname{dom}(A) = S_{n+1}$. If we take the family $A' = \{(x, y) \in A | x \in S_n\}$ [so $\operatorname{dom}(A') = S_n$ and $A'_i = \{(x, y) \in A' | x = i\}$] then $\bigcup_{i \in S_{n+1}} A_i = A_n \bigcup (\bigcup_{i \in S_n} A'_i)$

   Proof.

   a. If $x \in \bigcup_{i \in S_{n+1}} A_i$ then $\exists i \in \operatorname{dom}(A) = S_{n+1} \Rightarrow i < n+1$ such that $x \in A_i$ then we have the following cases

      i. $(\boldsymbol{i = n})$ then $x \in A_n \Rightarrow x \in A_n \bigcup (\bigcup_{i \in S_n} A'_i)$

      ii. $(\boldsymbol{i \neq n})$ then $i < n$ [if $n \leqslant i \underset{i \neq n}{\Rightarrow} n < i < n+1 \Rightarrow n+1 \leqslant i < n+1 \Rightarrow n+1 \neq n+1$ a contradiction] and thus $i \in S_n$. As $x \in A_i$ we have $(i, x) \in A \underset{i \in S_n}{\Rightarrow} (i, x) \in A' \Rightarrow x \in A'_i \Rightarrow x \in \bigcup_{i \in S_n} A'_i \Rightarrow x \in A_n \bigcup (\bigcup_{i \in S_n} A'_i)$

   in all cases we have $x \in A_n \bigcup (\bigcup_{i \in S_n} A'_i)$

b. If $x \in A_n \bigcup (\bigcup_{i \in S_n} A'_i)$ then we have the following cases

i. $\boldsymbol{(x \in A_n)}$ $\Rightarrow x \in \bigcup_{i \in S_{n+1}} A_i$

ii. $\boldsymbol{(x \notin A_n)}$ $\Rightarrow x \in \bigcup_{i \in S_n} A'_i \Rightarrow \exists i \in S_n \vdash x \in A'_i \Rightarrow (i, x) \in A' \subseteq A \Rightarrow x \in A_i \Rightarrow x \in \bigcup_{i \in S_{n+1}} A_i$

in all cases $x \in \bigcup_{i \in S_{n+1}} A_i$ □

As $n \in S$ we have $\bigcup_{i \in S_n} A'_i$ is finite and as by assumption A_n is finite we have by 5.37 that $\bigcup_{i \in S_{n+1}} A_i = A_n \bigcup (\bigcup_{i \in S_n} A'_i)$ is finite and thus $n + 1 \in S$.

By induction we have then $S = \mathbb{N}_0$ proving our theorem. □

Theorem 5.39. *If $\{A_i\}_{i \in I}$ is a family of finite sets with I finite then $\bigcup_{i \in I} A_i$ is finite*

Proof. As I is finite we have the existence of a bijection $f: n = S_n \rightarrow I$ $(n \in \mathbb{N}_0)$ and thus using 2.64 we have that $\bigcup_{i \in I} A_i = \bigcup_{j \in S_n} A_{f(j)}$ which is finite by the previous lemma. □

Theorem 5.40. *A set A is infinite if and only if $\exists B \subset A$ such that $B \approx A$ (if A is equipotent with a proper subset of itself)*

Proof.

First, suppose A is infinite then by 5.34 there exists a denumerable subset $B \subseteq A$. So there exists a bijection $b: \mathbb{N}_0 \rightarrow B$ define then the function $f: A \rightarrow A$ by $f(x) = \begin{cases} x \text{ if } x \in A \backslash B \\ b(b^{-1}(x) + 1) \ \forall x \in B \end{cases}$ then we have that

1. $\boldsymbol{(f(A) = A \backslash \{b(0)\})}$

Proof. If $y \in f(A)$ then there exists a $x \in A$ such that $y = f(x)$ and then either

a. $\boldsymbol{(x \in A \backslash B)}$ and thus $f(x) = x \in A \backslash B \Rightarrow f(x) = x \neq b(0) \in B \Rightarrow f(x) \in A \backslash \{b(0)\}$

b. $\boldsymbol{(x \in B)}$ and thus as $b^{-1}(x) \in \mathbb{N}_0 \Rightarrow b^{-1}(x) \geqslant 0 \Rightarrow b^{-1}(x) + 1 > 0$, if we would have $b(b^{-1}(x) + 1) = b(0)$ then as b is a bijection we have $0 = b^{-1}(x) + 1 > 0$ giving the contradiction $0 > 0$ so we must have that $f(x) = b(b^{-1}(x) + 1) \neq b(0) \Rightarrow f(x) \in A \backslash \{b(0)\}$ giving $f(A) \subseteq A \backslash \{b(0)\}$.

Also if $y \in A \backslash \{b(0)\}$ then we have either

a. $\boldsymbol{(y \notin B)}$ $\Rightarrow y \in (A \backslash \{b(0)\}) \backslash B$ then $f(y) = y \Rightarrow y \in f(A)$

b. $\boldsymbol{(y \in B)}$ then as b is a bijection $y = b(b^{-1}(y))$ and as $y \neq b(0)$ we can not have $b^{-1}(y) = 0$ [otherwise $y = b(b^{-1}(y)) = b(0)$], so $0 < b^{-1}(y) \Rightarrow 1 \leqslant b^{-1}(y) \underset{4.62}{\Rightarrow} \exists m \in \mathbb{N}_0 \vdash m + 1 = b^{-1}(y)$ and thus $b(m + 1) = b(b^{-1}(y)) = y$. Take now $x = b(m) \in B \subseteq A$ then $f(x) = b(b^{-1}(x) + 1) = b(m + 1) = y \Rightarrow y \in f(A)$.

giving that $A \backslash \{b(0)\} \subseteq f(A)$ and thus together with the already proven $f(A) \subseteq A \backslash \{b(0)\}$ that $f(A) = A \backslash \{b(0)\}$ □

2. **(f is injective)**

   Proof. If $f(x)=f(x')$ then we have the following possible cases to consider

   a. $\boldsymbol{(x,y\in A\backslash B)}$ then $x=f(x)=f(x')=x'\Rightarrow x=x'$

   b. $\boldsymbol{(x\in A\backslash B\wedge x'\in B)}$ then $f(x')=f(x)=x\in A\backslash B\Rightarrow f(x')\notin B$ contradicting the fact that $f(x')=b(b^{-1}(x')+1)\in B$, so this case does not apply.

   c. $\boldsymbol{(x\in B\wedge x'\in A\backslash B)}$ then $f(x)=f(x')=x'\in A\backslash B\Rightarrow f(x)\notin B$ contradicting the fact that $f(x)=b(b^{-1}(x)+1)\in B$ so this case does not apply.

   d. $\boldsymbol{(x,\ x'\in B)}$ then we have $b(b^{-1}(x)+1)=f(x)=f(x')=b(b^{-1}(x')+1)\underset{b\,\text{is injective}}{\Rightarrow} b^{-1}(x)+1=b^{-1}(x')+1\Rightarrow b^{-1}(x)=b^{-1}(x')\Rightarrow x=x'$

   so in all the cases that doe not lead to a contradiction we have $x=x'$ proving injectivity. □

From (1) and (2) we conclude that $f\colon A\rightarrow A\backslash\{b(0)\}$ is a bijection and thus that A is bijective with a proper subset of itself.

Second assume that there exists a proper subset $B\subset A$ and a bijection $f\colon A\rightarrow B$, giving a injective function $f\colon A\rightarrow A$ with $f(A)=B$. As $B\subset A$ there exists a $c\in A\backslash B$ and thus $c\notin f(A)$. Using 4.19 there exists a injective function $\lambda\colon\mathbb{N}_0\rightarrow A$ such that

1. $\lambda(0)=c$
2. $\forall n\in\mathbb{N}_0\vDash\lambda(s(n))=f(\lambda(n))$

So we found a bijective function $\lambda\colon\mathbb{N}_0\rightarrow\lambda(\mathbb{N}_0)\subseteq A$ giving a denumerable set $\lambda(\mathbb{N}_0)\subseteq A$ which by 5.34 means that A is infinite. □

Theorem 5.41. *If $n,m\in\mathbb{N}_0$ and $n\approx m$ then $n=m$*

Proof. Assume $n\approx m$ then

1. If $n<m\underset{n=S_n,m=S_m}{\Rightarrow}n\subset m$ and thus m is bijective to a proper subset of itself, which means by the previous theorem that m is infinite contradicting the finiteness of m (as it is bijective with itself)
2. If $m<n\underset{n=S_n,m=S_m}{\Rightarrow}m\subset n$ and thus n is bijective to a proper subset of itself, which means by the previous theorem that n is infinite contradicting the finiteness of n (as it is bijective with itself)

as $\mathbb{N}_0$ is totally ordered we must conclude that $n=m$. □

The previous theorem leads to the following observation: If A is a finite set then there exists a $n\in\mathbb{N}_0$ such that $n\approx A$, if there was also a $n'\in\mathbb{N}_0$ such that $n'\approx A$ and thus $n\approx n'$ and this leads to the following definition of the number of elements in a set.

Definition 5.42. *If A is a finite set then $\exists!n\in\mathbb{N}_0\vDash n\approx A$ this number is called the number of elements in A and is noted as $\#(A)$.*

We have already proved that a finite union of a finite sets is finite (see 5.39) but we can go further in case of a family of pairwise disjoint sets and calculate the number of elements in the union.

Theorem 5.43. *Let $\{A_i\}_{i\in\{1,\ldots,n\}}$ be such that $\forall i\in\{1,\ldots,n\}$ we have that A_i is finite, $\#(A_i)=n_i$ and $\forall i,j\in\mathbb{N}$ with $i\neq j$ we have $A_i\bigcap A_j=\emptyset$ then $\#\left(\bigcup_{i\in\{1,\ldots,n\}}A_i\right)=\sum_{i=1}^{n}\#(A_i)$*

Proof. We prove this by induction on n so let $\mathcal{S}=\left\{n\in\mathbb{N}|\text{if } \{A_i\}_{i\in\{1,\ldots,\}} \text{ is a family of pairwise disjoint finite sets then } \#\left(\bigcup_{i\in\{1,\ldots,n\}}A_i\right)=\sum_{i=1}^{n}\#(A_i)\right\}$ then we have

$1\in\mathcal{S}$. As $\#\left(\bigcup_{i\in\{1,\ldots,1\}}A_i\right)=\#(A_1)=\sum_{i=1}^{1}\#(A_i)$ we have $1\in\mathcal{S}$

$n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$. If $\{A_i\}_{i\in\{1,\ldots,n+1\}}$ is a pairwise disjoint family of sets then we have either

$A_{n+1}=\emptyset$. then $\#(A_{n+1})=0$ and $\#\left(\bigcup_{i\in\{1,\ldots,n+1\}}A_i\right)=\#\left(\bigcup_{i\in\{1,\ldots,n\}}A_i\right)\underset{n\in\mathcal{S}}{=}\sum_{i=1}^{n}\#(A_i)=\sum_{i=1}^{n+1}\#(A_i)$ proving $n+1\in\mathcal{S}$

$A_{n+1}\neq\emptyset$. Define then $N=\sum_{i=1}^{n}\#(A_i)$ then we have as as $n\in\mathcal{S}$ the existence of a bijection $\sigma_1\colon\{1,\ldots,N\}\to\bigcup_{i=1}^{n}A_i$ and as A_{n+1} is finite there exists a bijection $\sigma_2\colon\{1,\ldots,\#(A_{n+1})\}\to A_{n+1}$. Define $\sigma\colon\{1,\ldots,N+\#(A_{n+1})\}\to\bigcup_{i\in\{1,\ldots,n+1\}}A_i$ by $i\to\sigma(i)=\begin{cases}\sigma_1(i) \text{ if } i\in\{1,\ldots,N\}\\ \sigma_2(i-N) \text{ if } i\in\{N+1,\ldots,N+\#(A_{n+1})\}\end{cases}$, then we have

σ is injective. Take $i,j\in\{1,\ldots,N+\#(A_{n+1})\}$ such that $\sigma(i)=\sigma(j)$ then we have either

$i,j\in\{1,\ldots,N\}$. then $\sigma_1(i)=\sigma(i)=\sigma(j)=\sigma_1(j)$ proving $i=j$

$i,j\in\{N+1,\ldots,N+\#(A_{n+1})\}$. then $\sigma_2(i-N)=\sigma(i)=\sigma(j)=\sigma_2(j-N)$ proving $i-N=j-N\Rightarrow i=j$

$i\in\{1,\ldots,N\}\wedge j\in\{N+1,\ldots,N+\#(A_{n+1})\}$. then $\sigma(i)=\sigma_1(i)\in\bigcup_{i\in\{1,\ldots,n\}}A_i$ and $\sigma(i)=\sigma(j)=\sigma_2(j-N)\in A_{n+1}$ proving that $\sigma(i)\in\left(\bigcup_{i\in\{1,\ldots,n\}}A_i\right)\bigcap A_{n+1}=\emptyset$ a contradiction so this case does not occur.

$j\in\{1,\ldots,N\}\wedge i\in\{N+1,\ldots,N+\#(A_{n+1})\}$. then $\sigma(j)=\sigma_j(j)\in\bigcup_{i\in\{1,\ldots,n\}}A_i$ and $\sigma(j)=\sigma(i)=\sigma_2(i-N)\in A_{n+1}$ proving that $\sigma(j)\in\left(\bigcup_{i\in\{1,\ldots,n\}}A_i\right)\bigcap A_{n+1}=\emptyset$ a contradiction so this case does not occur.

σ is surjective. If $y\in\bigcup_{i\in\{1,\ldots,n+1\}}A_i$ then we have either

$y\in A_{n+1}$. then as σ_2 is a bijection there exists a $i\in\{1,\ldots,\#(A_{n+1})\}$ such that $\sigma_2(j)=y$. Take $j=i+N\in\{N+1,\ldots,N+\#(A_{n+1})\}$ then $\sigma(j)=\sigma_2(j-N)=\sigma_2(i)=y$

$y\in\bigcup_{i\in\{1,\ldots,n\}}A_i$. then as σ_1 is a bijection there exists a $i\in\{1,\ldots,\sum_{i=1}^{N}\#(A_i)\}$ such that $\sigma_1(i)=y\Rightarrow\sigma(i)=\sigma_1(i)=y$

As σ is a bijection we have that $\#\left(\bigcup_{i\in\{1,\ldots,n+1\}} A_i\right) = N + \#(A_i) = \sum_{i=1}^{N} \#(A_i) + \#(A_{n+1}) = \sum_{i=1}^{N+1} \#(A_i)$ proving $N+1 \in \mathcal{S}$ □

Theorem 5.44. *If A, B are finite sets with $n = \#(A)$ and $m = \#(B)$ then $A \times B$ is finite and has size $n \cdot m$*

Proof. If $A = \emptyset$ (or $B = \emptyset$) then $\#(A) = 0$ (or $\#(B) = 0$) and then $A \times B = \emptyset$ and $\#(A \times B) = 0 = \#(A) \cdot \#(B)$. So assume that $A, B \neq \emptyset$ then there exists bijections $a: B \rightarrow S_n$, $b: B \rightarrow S_m$ where $n, m \in \mathbb{N}$. Define now $f: A \times B \rightarrow S_{m \cdot n}$ by $(x, y) \rightarrow a(x) \cdot m + b(y)$ (as $\forall x \in A, \forall y \in B$ we have $0 \leqslant a(x) < n$ and $0 \leqslant b(y) < m$ we have indeed that $\forall (x, y) \in A \times B$ that $0 \leqslant f(x, y) < n \cdot m + m$). We prove now that f is a bijection:

1. **(injectivity)** If $f(x, y) = f(x', y')$ then $a(x) \cdot m + b(y) = a(x') \cdot m + b(y')$ and using the fact that $b(y), b(y') < m$ and 4.74 we have that $a(x) = a(x')$ and $b(y) = b(y') \underset{a, b \text{are bijections}}{\Rightarrow} x = x'$ and $y = y' \Rightarrow (x, y) = (x', y')$ proving injectivity.
2. **(surjectivity)** If $z \in S_{m \cdot n}$ then $0 \leqslant z < m \cdot n$, using the Division Algorithm (see 6.48) there exists a q, r so that $z = q \cdot m + r$ and $0 \leqslant r < m \Rightarrow r \in S_m$, also $0 \leqslant q < n$ [if $n \leqslant q \Rightarrow n \cdot m \leqslant q \cdot m \Rightarrow n \cdot m + r \leqslant q \cdot m + r = z \Rightarrow n \cdot m + r \leqslant z \Rightarrow n \cdot m \leqslant z < n \cdot m$ a contradiction] and thus $q \in S_n$. Now as a, b are bijections there exists $x \in A$ and $y \in B$ such that $a(x) = n$ and $b(x) = m$ and thus $f(x, y) = a(x) \cdot m + b(y) = q \cdot m + r = z$. □

Theorem 5.45. *If A is a finite set and $B \subseteq A$ then $B, A \backslash B$ are finite and $\#(B) \leqslant \#(A)$ and $\#(A \backslash B) = \#(A) - \#(B)$. Note that from the last equation it follows that $\#(A \backslash B) + \#(B) = \#(A)$ and thus that $\#(A \backslash B) \leqslant \#(B)$*

Proof. As A is finite with $n = \#(A)$ then there exists a bijection $a: n = S_n \rightarrow A$ we have now to consider the following cases:

1. $(\boldsymbol{B = A})$ then obviously B is finite and $A \backslash B = \emptyset$ and thus $A \backslash B$ is finite and $\#(A \backslash B) = \#(A) - \#(B) = \#(A) - \#(A) = 0$
2. $(\boldsymbol{B = \emptyset})$ then obviously B is finite and $A \backslash B = A$ and thus $A \backslash B$ is finite and $\#(A \backslash B) = \#(A) = \#(A) - 0 = \#(A) - \#(B)$
3. $(\boldsymbol{\emptyset \neq B \subset A})$ Let now $X = \{n \in \{2, \ldots, \} |$ If A is such that $\#(A) = n$ and $\emptyset \neq B \subset A$ then there exists a bijection $f: S_n \rightarrow A$ and a $m \in S_n$ such that $f_{|S_m}: S_m \rightarrow B$ is a bijection$\}$ then we have
   a. If $n = 2$ then if $\#(A) = 2$ then $2 = \{0, 1\} = S_2$ and there exists a bijection $f': \{0, 1\} \rightarrow A$, now as $\emptyset \neq B \subset A$ there exists a $a \in A \backslash B$ and as f' is a bijection then there exists a $i \in \{0, 1\}$ such that $f'(i) = a$ we have now two cases:
      i. $(\boldsymbol{i = 0})$ then $B = \{f'(1)\}$ so take the bijection $f: \{0, 1\} \rightarrow A$ with $\left\{ \begin{array}{l} f(0) = f'(1) \\ f(1) = f'(0) \end{array} \right.$ then $f_{|S_1 = \{0\}}: \{0\} \rightarrow \{f'(1)\} = B$ is a bijection
      ii. $(\boldsymbol{i = 1})$ then $B - \{f'(0)\}$ and take then $f = f'$ so that $f_{|S_1} = f'_{|S_1}: \{0\} \rightarrow \{f'()\} = B$ is a bijection

we can thus conclude that $2\in X$

b. Assume now that $n\in X$. If now $\#(A)=n+1$ then there exists a bijection $f'\colon S_{n+1}\to A$, now as $\emptyset\neq B\subset A$ there exists a $a\in A\backslash B$ and thus $\emptyset\neq B\subseteq A\backslash\{a\}$. As f' is a bijection there exists a $i\in S_{n+1}=\{0,\ldots,n\}$ such that $f'(i)=a$. We have now two cases to consider:

   i. $\boldsymbol{(i=n)}$ take then $f=f'$

   ii. $\boldsymbol{(i\neq n)}$ take then $a'=f'(n)$ then $a\neq a'$[otherwise $i=n$]. Then $f'_{|S_{n+1}\backslash\{i,n\}}\colon S_{n+1}\backslash\{i,n+1\}\to A\backslash\{a,a'\}$ is a bijection (see 2.51). Define further the bijection $f''\colon\{i,n\}\to\{a,a'\}$ by $\left\{\begin{smallmatrix} f''(i)=a' \\ f''(n)=a\end{smallmatrix}\right.$ so that we have a bijection (see 2.43) $f=f'_{|S_{n+1}\backslash\{i,n\}}\bigcup f''\colon S_{n+1}\to A$ with $(n)=f''(n)=a$.

So we have found a bijection $f\colon S_{n+1}\to A$ such that $f(n)=a$. We have now the following cases to consider for $A\backslash\{a\}$ (remember $B\subseteq A\backslash\{a\}$):

   i. $\boldsymbol{(\emptyset\neq B=A\backslash\{\alpha\})}$ then we take $m=n$ and $f_{|S_n}\colon S_n\to A\backslash\{a\}=B$ is a bijection (as $f(\{n\})=\{a\}$ and thus $n+1\in X$

   ii. $\boldsymbol{(\emptyset\neq B\subset A\backslash\{a\})}$ then as $n=\#(A\backslash\{a\})$ and $n\in X$ there exists a bijection $f'''\colon S_n\to A\backslash\{a\}$ and a $m\in S_n$ such that $f'''_{|S_m}\colon S_m\to B$ is a bijection, take then the bijection $f^{\text{iv}}\colon\{n\}\to\{a\}$ by $f^{\text{iv}}(n)=a$ and construct then the bijection $f^v=f'''\bigcup f^{\text{iv}}\colon S_{n+1}\to A$. We have then that $f^v_{|S_m}=f''_{|S_m}$ is a bijection from S_m to B proving that $n+1\in X$

Using mathematical induction (see 4.77) we have that $X=\{2,\ldots\}$, now as $\emptyset\neq B\subset A$ we must have $2\leqslant\#(A)=n\Rightarrow n=\#(A)\in\{2,\ldots\}=X$ so there exists a bijection $f\colon S_n\to A$ and a $m\in S_n$ such that $f_{|S_m}\colon S_m\to B$ is a bijection, this proves that B is finite and $\#(B)=m$. Now if $i\in\{m,\ldots,n-1\}$ we have $f(i)\notin B$ [if $f(i)\in B$ then as $f_{|S_m}$ is a bijection there exists a $j<m$ such that $f(j)=f_{|S_m}(j)=f(i)\Rightarrow i=j<m$ contradicting $i\in\{m,\ldots,n-1\}$] and thus $f(\{m,\ldots,n-1\})\subseteq A\backslash B$. If $y\in A\backslash B$ then as f is a bijection there exists a $i\in S_n$ such that $y=f(i)$ if now $i<m\Rightarrow i\in S_m\Rightarrow f(i)\in B$ a contradiction so we must have $i\in\{m,\ldots,n-1\}\Rightarrow A\backslash B\subseteq f(\{m,\ldots,n-1\})$. So $f(\{m,\ldots,n-1\})=A\backslash B$ and thus by 2.51 we have that $f_{|\{m,\ldots,n-1\}}\colon\{m,\ldots,n-1\}\to A\backslash B$ is a bijection. Define now $g\colon S_{n-m}\to\{m,\ldots,n-1\}$ (note that from $m\in S_n\Rightarrow m<n\Rightarrow 0<n-m\in\mathbb{N}_0$) by $i\to g(i)=i+m$ then g is a bijection:

a. **(injectivity)** if $g(i)=g(j)\Rightarrow i+m=j+m\Rightarrow i=j$

b. **(surjectivity)** If $j\in\{m,\ldots,n-1\}\Rightarrow m\leqslant j<n-1\Rightarrow 0\leqslant j-m<n-m-1\Rightarrow j-m\in S_{n-m}$ and $g(j-m)=j$

So we have then a bijection $f_{|\{m,\ldots,n-1\}}\circ g\colon S_{n-m}\to A\backslash B$ proving that $A\backslash B$ is finite and $\#(A\backslash B)=n-m=\#(A)-\#(B)$ □

Theorem 5.46. *If A is a finite set then if $B\subset A$ we have that $\#(B)<\#(A)$*

Proof. Using the previous theorem we have that B is finite and $\#(B) \leqslant \#(A)$, if now $\#(A) = \#(B) = n$ then $A \approx n \approx B$ and thus $A \approx B$ which by 5.40 would give that A is infinite contradicting the finiteness of A. So we must have $\#(B) < \#(A)$ □

Theorem 5.47. *If A is a set, $n \in \mathbb{N}_0$ and $f: S_n \rightarrow A$ a surjection then A is finite and $\#(A) \leqslant n$.*

Proof. If $n = 0$ then $S_0 = \emptyset$ and if $f: \emptyset \rightarrow A$ is a surjection then if $x \in A$ there exists a $i \in \emptyset$ such that $x = f(i)$ which would mean that $\emptyset \neq \emptyset$ a contradiction, so in this cases we must have $A = \emptyset \Rightarrow A$ is finite and $\#(A) = 0 \leqslant 0$ proving the theorem for the case $n = 0$. We prove the case $n \in \mathbb{N}$ by mathematical induction so let $X = \{n \in \{1, ...\} |$ if A is a set such that there exists a surjection $f: S_n \rightarrow A$ then A is finite and $\#(A) = n\}$ we have then that

1. If $n = 1$ then if A is such that there exists a surjection $f: S_1 = \{0\} \rightarrow A$ then $A = \{f(0)\}$ which trivially means that A is finite and that $\#(A) = 1 \leqslant 1 \Rightarrow 1 \in X$

2. If now $n \in X$ then if A is such that there exists a surjection $f: S_{n+1} \rightarrow A$. Then as $f(n) \in A$ we can not have $A = \emptyset$ and we are left with the following two cases

   a. $(\forall \boldsymbol{a} \in \boldsymbol{A} \vDash \boldsymbol{a} = \boldsymbol{f}(\boldsymbol{n}))$ $\Rightarrow A = \{f(n)\}$ and thus A is finite with $\#(A) = 1 \leqslant n+1$. This gives $n+1 \in X$

   b. $(\exists \boldsymbol{a} \in \boldsymbol{A} \vdash \boldsymbol{a} \neq \boldsymbol{f}(\boldsymbol{n}))$ Define then $g: S_n \rightarrow A \backslash \{f(n)\}$ by

$$\begin{aligned} g(i) &= f(i) \quad \text{if } i \in S_n \backslash f^{-1}(\{f(n)\}) \, [\Rightarrow g(i) \in A \backslash \{f(n)\}] \\ &= a \quad\ \ \text{if } i \in S_n \bigcap f^{-1}(\{f(n)\}) \, [\Rightarrow g(i) \in A \backslash \{f(n)\}] \end{aligned}$$

   If now $y \in A \backslash \{f(n)\}$ then as f is a surjection there exists a $i \in S_{n+1}$ such that $f(i) = y$. We can not have $i = n$ [As then $f(i) = f(n) \Rightarrow y \in \{f(n)\}$], also we can not have $i \in f^{-1}(\{f(n)\})$ [as then $f(i) \in \{f(n)\} \Rightarrow y \in \{f(n)\}$ so $i \in S_n \backslash f^{-1}(\{f(n)\})$ and thus $g(i) = f(i) = y$. This proves that g is a surjection. As $n \in X$ we have that $A \backslash \{(f(n))\}$ is finite and $m = \#(A \backslash \{f(n)\}) \leqslant n$. So there exists a bijection $h: S_m \rightarrow A \backslash \{f(n)\}$, combine this with the bijection $h': \{m\} \rightarrow \{f(n)\}$ using 2.43 to form the bijection $h \bigcup h': S_{m+1} \rightarrow A$. So A is finite and $\#(A) = m+1 \leqslant n+1$ and this proves that $n+1 \in X$

Using mathematical induction (4.77) we have that $X = \{1, ...\} = \mathbb{N}$ proving the theorem for $n \in \mathbb{N}$ □

Corollary 5.48. *If A, B are sets where A is finite and $f: A \rightarrow B$ is a surjection then B is finite and $\#(B) \leqslant \#(A)$*

Proof. If A is finite and $\#(A) = n$ then there exists a bijection $b: n \rightarrow A$ so $f \circ g$: $S_n \rightarrow B$ is a surjection and by the previous theorem we have then B is finite and $\#(B) \leqslant n = \#(A)$ □

Theorem 5.49. *Let A, B be sets, A infinite and $f: A \rightarrow B$ a injection then B is infinite*

Proof. Assume that B is finite then $f(A) \subseteq B$ is finite and there is a bijection $f: \{1, \ldots, n\} \rightarrow f(A)$ so that as $f: A \rightarrow f(A)$ is a bijection we find that $f^{-1} \circ h$: $\{1, \ldots, n\} \rightarrow A$ is a bijection making A finite a contradiction. □

Theorem 5.50. *Let $\langle A, \leqslant \rangle$ be a fully ordered non empty finite set then* $\max(A)$ *(or* $\min(A)$*) exists (see 2.163)*

Proof. We prove this by induction on $\#(A) = n$. So let $S = \{n \in \mathbb{N} = \{1, \ldots\} | \text{if } \#(A) = n \text{ then } \max(A) \text{ exists}\}$ then we have :

1. If $n = 1$ then $A = \{a\}$ with $\max(A)$
2. If $n \in S$ let then $\#(A) = n + 1$ so there exists a bijection $b: \{1, \ldots, n + 1\} \rightarrow A$, take then $A \setminus \{b(n + 1)\}$ which is bijective to $\{1, \ldots, n\}$ so that $\#(A \setminus \{b(n + 1)\}) = n$ and thus $m' = \max(A \setminus \{b(n + 1)\})$ exist. We have now two cases to consider:
   a. $\boldsymbol{(m' \leqslant b(n + 1))}$ then $\max(A) = b(n + 1)$
   b. $\boldsymbol{(b(n + 1) \leqslant m')}$ then $\max(A) = m'$

The proof for the minimum is similar. □

Theorem 5.51. *If A is finite and $f: \mathbb{N} \rightarrow A$ is a function then $\exists a \in A$ such that $\{m \in \mathbb{N} | f(x) = a\} = f^{-1}(\{a\})$ is infinite.*

Proof. Assume that the theorem is not valid then $\forall a \in A$ we have that $f^{-1}(\{a\})$ is finite so that $\bigcup_{a \in A} f^{-1}(\{a\})$, being a finite union of finite sets, is finite. Now from $\mathbb{N} = f^{-1}(A) = f^{-1}(\bigcup_{a \in A} \{a\}) = \bigcup_{a \in A} f^{-1}(\{a\})$ we would then conclude that $\mathbb{N}$ is finite which is a contradiction. □

Corollary 5.52. *If A is finite and $f: \mathbb{N} \rightarrow A$ is a function then $\exists a \in A$ so that $\forall n \in \mathbb{N}$ there exists a $m \in \{i \in \mathbb{N} | i \geqslant n\}$ so that $f(m) = a$*

Proof. By the previous theorem $\exists a \in A$ such that $f^{-1}(\{a\}) = \{m \in \mathbb{N} | f(m) = a\}$ is infinite. We proceed non by contradiction. So assume that $\exists n \in \mathbb{N}$ such that $\forall m \geqslant n$ we have $f(m) \neq a$. If then $m \in f^{-1}(\{m\}) \Rightarrow f(m) = a \Rightarrow m < n \Rightarrow f^{-1}(\{a\}) \subseteq S_n$ meaning that $f^{-1}(\{a\})$ is finite contradicting the fact that $f^{-1}(\{a\})$ is infinite. □

Theorem 5.53. *Let $\langle X, \leqslant \rangle$ be a fully ordered set and let $A \subseteq X$ be a non empty finite set. Then there exists a bijection $i: \{1, \ldots, \#(A)\} \rightarrow A$ such that $\forall k \in \{1, \ldots, \#(A) - 1\}$ we have $i(k) < i(k + 1)$. Second for such a bijection we must have $\forall k, l \in \{1, \ldots, \#(A)\}$ with $k < l$ that $i(k) < i(l)$. Third the above bijection is unique.*

Proof. If $\#(A) = 1$ we take the bijection $i = 1_{\{1\}}$ then for $k \in \{1, \ldots, 0\} = \emptyset$ we have that $i(l) < i(k)$ is satisfied vacuously, also $\forall k, l \in \{1, \ldots, 1\}$ with $k < l$ we have that $i(k) < i(l)$ is satisfied vacuously. So we only have to prove the theorem for $\#(A) > 1$. We prove the rest by induction so let $\mathcal{S} = \{n \in \{2, \ldots\} | \text{if } \#(A) = n \text{ then there exists a bijection } i: \{1, \ldots, n\} \rightarrow A \text{ with } \forall k \in \{1, \ldots, n - 1\} \text{ we have } i(k) < i(k + 1)\}$. We have then:

1. If $n = 2$ then there exists a bijection $b: \{1, \ldots, 2\} = \{1, 2\} \rightarrow A$ so $A = \{b(1), b(2)\}$. If $b(1) < b(2)$ we have found our bijection. If $b(2) < b(1)$ construct then $i: \{1, 2\} \rightarrow A$ by $i(1) = b(2)$ and $i(2) = b(1)$ which is obviously a bijection with $i(1) < i(2)$. This proves that $2 \in \mathcal{S}$.

2. Assume that $n \in \mathcal{S}$ and that we have a finite A with $\#(A) = n + 1$. As A is finite a maximum exists (see 5.50) take then $A \setminus \{\max(A)\}$. Using 5.45 we have that $\#(A \setminus \{\max(A)\}) = \#(A) - \#(\{\max(A)\}) = (n + 1) - 1 = n$. As $n \in \mathcal{S}$ there exists a bijection i': $\{1, \ldots, n\} \rightarrow A \setminus \{\max(A)\}$ such that $\forall k \in \{1, \ldots, n - 1\}$ we have $'(k) < i'(k + 1)$. Define then the bijection i: $\{1, \ldots, n + 1\} \rightarrow A$ by $k \rightarrow i(k) = \left\{ \begin{smallmatrix} i'(k) \text{ if } k \in \{1, \ldots, n\} \\ \max(A) \text{ if } k = n + 1 \end{smallmatrix} \right.$ (see 2.43) then if $k \in \{1, \ldots, (n + 1) - 1\} = \{1, \ldots, n\}$ we have either $k \in \{1, \ldots, n - 1\}$ and then $i(k) = i'(k) < i'(k + 1) = i(k)$, if $k = n$ then $i(k) \leqslant \max(A) = i(k + 1)$. As $i(k) = i'(k) \in A \setminus \{\max(A)\}$ we must have $i(k) \neq \max(A) = i(k + 1)$ so that we have $i(k) < i(k + 1)$. This proves that $n + 1 \in \mathcal{S}$

Using induction we have $\mathcal{S} = \{2, \ldots\}$ proving the existence. Next we prove by induction that if i: $\{1, \ldots, \#(A)\} \rightarrow A$ is a bijection such that $\forall k \in \{1, \ldots, \#(A) - 1\}$ we have $i(k) < i(k + 1)$ then we have $\forall k, l \in \{1, \ldots, \#(A)\}$ that $i(k) < i(l)$. So let $k \in \{1, \ldots, \#(A)\}$ and take $\mathcal{P}_k = \{m \in \mathbb{N} | \text{if } k + m \leqslant \#(A) \text{ then } i(k) < i(k + m)\}$ then we have

1. If $m = 1$ then if $k + 1 < \#(A)$ we have $k \leqslant \#(A) - 1$ so that $i(k) < i(k + 1) = i(k + m)$ proving that $1 \in \mathcal{P}_k$
2. Let $m \in \mathcal{P}_k$ then if $k + (m + 1) \leqslant \#(A)$ we have $(k + m) \leqslant \#(A) - 1$ so that $i(k + m) < i((k + m) + 1) = i(k + (m + 1))$ and as $m \in \mathcal{P}_k$ and $k + m \leqslant \#(A)$ we have $i(k) < i(k + m)$ so that $i(k) < i(k + (m + 1))$ proving that $m + 1 \in \mathcal{P}_k$

Using induction we have that $\mathcal{P}_k = \mathbb{N}$ so if $k, l \in \{1, \ldots, \#(A)\}$ and $k < l$ then $m = l - k \in \mathbb{N} = \mathcal{P}_k$ and as $k + m = l \leqslant \#(A)$ we have that $i(k) < i(k + m) = i(l)$.

Finally to prove uniqueness suppose that $n = \#(A)$ and i_1: $\{1, \ldots, n\} \rightarrow A$, i_2: $\{1, \ldots, n\} \rightarrow A$ are bijections such that if $k \in \{1, \ldots, n - 1\}$ then $i_1(k) < i_1(k) \wedge i_2(k) < i_2(k + 1)$. Take now $\mathcal{Q} = \{k \in \mathbb{N} | \text{if } k \leqslant n \text{ then } \forall l \in \{1, \ldots, k\} \text{ we have } i_1(l) = i_2(l)\}$ then we have:

1. If $k = 1$ then if $i_1(1) \neq i_2(1)$ we have either
   a. $i_1(1) < i_2(1)$ then as $i_1(1) \in A$ there exists a $l \in \{1, \ldots, n\}$ such that $i_2(l) = i_1(1)$. If now $l > 1$ we must have $i_2(1) < i_2(l) = i_1(1) < i_2(1)$ a contradiction, so we must have $l = 1$ but this means that $i_2(1) = i_1(1)$ contradicting $i_1(1) < i_2(1)$. So this case leads to a contradiction.
   b. $i_2(1) < i_1(1)$ then as $i_2(1) \in A$ there exists a $l \in \{1, \ldots, n\}$ such that $i_1(l) = i_2(1)$. If now $l > 1$ we must have $i_1(1) < i_1(l) = i_2(1) < i_1(1)$ a contradiction, so we must have $l = 1$ but this means that $i_1(1) = i_2(1)$ contradicting $i_2(1) < i_1(1)$. So this case leads to a contradiction.

   we must thus conclude that $i_1(1) = i_2(1)$ proving that $1 \in \mathcal{Q}$
2. Let $k \in \mathcal{Q}$ then if $k + 1 \leqslant n$ we prove by contradiction that $i_1(k + 1) = i_2(k + 1)$. So assume that $i_1(k + 1) \neq i_2(k + 1)$ then we have either
   a. $i_1(k + 1) < i_2(k + 1)$ then as $i_1(k + 1) \in A$ there exists a $l \in \{1, \ldots, n\}$ such that $i_2(l) = i_1(k + 1)$. If now $l = k + 1$ we would have the contradiction $i_1(k + 1) = i_2(l) = i_2(k + 1) > i_1(k + 1)$, if $l > k + 1$ then $i_1(k + 1) = i_2(l) > i_2(k + 1) > i_1(k + 1)$ again a contradiction, if $l < k + 1$ then $l \leqslant k$ so that $i_1(l) = i_2(l) = i_1(k + 1) > i_1(l)$ again a contradiction. So all the cases leads to a contradiction.

b. $i_2(k+1) < i_1(k+1)$ then as $i_2(k+1) \in A$ there exists a $l \in \{1, \ldots, n\}$ such that $i_1(l) = i_2(k+1)$. If now $l = k+1$ we would have the contradiction $i_2(k+1) = i_1(l) = i_1(k+1) > i_2(k+1)$, if $l > k+1$ then $i_2(k+1) = i_1(l) > i_1(k+1) > i_2(k+1)$ again a contradiction, if $l < k+1$ then $l \leqslant k$ so that $i_2(l) = i_1(l) = i_2(k+1) > i_2(l)$ again a contradiction. So all the cases leads to a contradiction.

so we must have $i_1(k+1) = i_2(k+1)$ or $k+1 \in \mathcal{Q}$

By induction we have thus that $\mathcal{Q} = \mathbb{N}$. So if $k \in \{1, \ldots, n\}$ then $k \in \mathcal{Q}$ and $k \leqslant n$ so that $i_1(k) = i_2(k)$ proving that $i_1 = i_2$. □

5.3 Properties of denumerable sets

Lemma 5.54. *Every subset of $\mathbb{N}_0$ is finite or denumerable.*

Proof. As $\langle \mathbb{N}_0, \leqslant \rangle$ is well-ordered (see 4.52) we can use 2.196. So if $A \subseteq \mathbb{N}_0$ we have either that A is isomorphic with $\mathbb{N}_0$ (and thus denumerable as a isomorphism is a bijection) or A is isomorphic with a segment $S_n = n$ and thus finite. □

Theorem 5.55. *Every subset of a denumerable set is finite or denumerable.*

Proof. Let A be a denumerable set and take $B \subseteq A$. From the denumerability of A there exists a bijective function $f: \mathbb{N}_0 \rightarrow A$. Then using the previous theorem we have that $f^{-1}(B) \subseteq \mathbb{N}_0$ is either denumerable or finite, so we have the following cases:

1. **($f^{-1}(B)$ is enumerable)** Then there exists a bijection $g: \mathbb{N}_0 \rightarrow f^{-1}(B)$, together with the bijection $f_{|f^{-1}(B)}: f^{-1}(B) \rightarrow B$ this means that B is bijective with $\mathbb{N}_0$ and is thus denumerable.
2. **($f^{-1}(B)$ is finite)** Then there exists a $n \in \mathbb{N}_0$ such that there exists a bijection $g: n \rightarrow f^{-1}(B)$, together with the bijection bijection $f_{|f^{-1}(B)}: f^{-1}(B) \rightarrow B$ this means that B is bijective with n and is thus finite.

□

Theorem 5.56. *If $0 < m$ then there exists a number noted by $m-1$ such that*

1. $(m-1)+1 = m$
2. $m-1 = m'-1 \Leftrightarrow m = m'$
3. $(m+1)-1 = m$

Note 5.57. As $0 < m \Rightarrow 1 \leqslant m$ we have that this definition is equivalent with the definition 4.63.

Proof.

1. If $0 < m$ then $1 \leqslant m$ and thus by 4.62 there exists a $k \in \mathbb{N}_0$ such that $1+k = k+1 = m$ so we can take $m-1$ to be k.
2. If $m-1 = m'-1 \Rightarrow m = (m-1)+1 = (m'-1)+1 = m' \Rightarrow m = m'$. If $m = m'$ then $(m-1)+1 = m = m' = (m'-1)+1 \Rightarrow s(m-1) = s(m'-1) \underset{4.15}{\Rightarrow} m-1 = m'-1$
3. We have $((m+1)-1)+1 = m+1 \Rightarrow s((m+1)-1) = s(m) \underset{4.15}{\Rightarrow} (m+1)-1 = m$ □

Theorem 5.58. $\mathbb{N}_0 \times \mathbb{N}_0 \approx \mathbb{N}_0$

Proof. First define the function $f: \mathbb{N}_0 \times \mathbb{N}_0 \rightarrow \mathbb{N}_0 \times \mathbb{N}_0$ (using the previous theorem to define $m-1$)

$$f(k,m) = \begin{cases} (0,k+1) \text{ if } m=0 \\ (k+1,m-1) \text{ if } m \neq 0 \ (\Rightarrow 0 < m) \end{cases}$$

then if $f(k,m) = f(k',m')$ we have the following cases for m, m'

1. $\boldsymbol{(m,m'=0)}$ then $(0,k+1) = f(k,m) = f(k',m') = (0,k'+1) \Rightarrow k+1 = k'+1 \Rightarrow k = k' \underset{m=0=m'}{\Rightarrow} (k,m) = (k',m')$
2. $\boldsymbol{(m>0,m'=0)}$ then $(k+1,m-1) = f(k,m) = f(k',m') = (0,k'+1) \Rightarrow k+1 = 0 \underset{0 \leqslant k \Rightarrow 0 < k+1}{\Rightarrow} 0 < 0$ a contradiction, this case can not occur
3. $\boldsymbol{(m=0,m'>0)}$ then $(0,k+1) = f(k,m) = f(k',m') = (k'+1,m'-1) \Rightarrow 0 = k'+1 \underset{0 \leqslant k' \Rightarrow 0 < k'+1}{\Rightarrow} 0 > 0$ a contradiction, this case can not occur
4. $\boldsymbol{(m>0,m'>0)}$ then $(k+1,m-1) = f(k,m) = f(k',m') = (k'+1,m'-1) \Rightarrow k+1 = k'+1 \wedge m-1 = m'-1 \Rightarrow k = k' \wedge m = m' \Rightarrow (k,m) = (k',m')$

this proves that $f: \mathbb{N}_0 \times \mathbb{N}_0 \rightarrow \mathbb{N}_0 \times \mathbb{N}_0$ is a injective function. If now $f(k,m) = (0,0)$ then we have the following cases to consider for m

1. $\boldsymbol{(m=0)}$ then $(0,0) = f(k,m) = (0,k+1) \Rightarrow 0 = k+1 \underset{0 \leqslant k \Rightarrow 0 < k+1}{\Rightarrow} 0 < 0$ a contradiction.
2. $\boldsymbol{(m \neq 0)}$ then $(0,0) = f(k,m) = (k+1,m-1) \Rightarrow 0 = k+1 \underset{0 \leqslant k \Rightarrow 0 < k+1}{\Rightarrow} 0 < 0$ a contradiction.

so we have $f(k,m) \neq (0,0)$ or $(0,0) \notin f(\mathbb{N}_0 \times \mathbb{N}_0)$.

Using 4.19 there exists a injective function $\lambda: \mathbb{N}_0 \rightarrow \mathbb{N}_0 \times \mathbb{N}_0$ such that

1. $\lambda(0) = (0,0)$
2. $\forall n \in \mathbb{N}_0 \vDash \lambda(n+1) = f(\lambda(n))$

We prove now the following **property (a)** of λ, assume that there exists a n, $m \in \mathbb{N}_0$ such that $\lambda(n) = (0,m)$ then $\forall k \in \mathbb{N}_0$ with $\exists l \in \mathbb{N}_0$ such that $k+l = m$ we have $\lambda(n+k) = (k,l)$. The proof is by induction on k so let $S_{n,m} = \{k \in \mathbb{N}_0 | \text{if } \exists l \in \mathbb{N}_0 \text{ such that } k+l = m \text{ then } \lambda(n+k) = (k,l)\}$ then:

1. If $k=0$ then if $l=m$ we have $k+l=m$ and $\lambda(n+k) = \lambda(n) = (0,m) = (k,l)$ so we have $0 \in S_{n,m}$
2. Assume $k \in S_{n,m}$ then consider $k+1$ if there exists a $l \in \mathbb{N}_0$ such that $(k+1)+l = m$ then $k+(l+1) = m$ and as $k \in S_{n,m}$ we have $\lambda(n+k) = (k,l+1)$, now as $0 \leqslant l \Rightarrow 0 < l+1$ we have $\lambda(n+(k+1)) = \lambda((n+k)+1) = f(\lambda(n+k)) = f(k,l+1) = (k+1,l)$ so $k+1 \in S_{n,m}$

Thus if $k+l = m \Rightarrow k \in \mathbb{N}_0 = S_{n,m} \Rightarrow \lambda(n+k) = (k,l)$

We shown now by induction that λ is surjective, so let $S = \{n \in \mathbb{N}_0 | \forall (k,m) \in \mathbb{N}_0 \times \mathbb{N}_0 \vdash k+m = n \vDash \exists n' \in \mathbb{N}_0 \vdash \lambda(n') = (k,m)\}$.

1. If $n=0$ take $k=m \in \mathbb{N}_0$ is such that $k+m=0$ then we must have $k=0=m$ [if $m>0$ then $0=k+m>0 \Rightarrow 0 \neq 0$ and if $k>0$ again $0=k+m>0 \Rightarrow 0 \neq 0$] so that $\lambda(0) = (0,0) = (k,m)$ and thus $0 \in S$.

2. Assume now that $n \in S$ and consider $n+1$. If $(k, m) \in \mathbb{N}_0 \times \mathbb{N}_0$ is such that $k+m = n+1$. we can have then the following cases for k

   a. **($k = 0$)** then $k = k+m = n+1$ and thus $(k, m) = (0, m) = (0, n+1) = f(n, 0)$. As $n+0 = n \in S$ there exists a $n'' \in \mathbb{N}_0$ such that $(n, 0) = \lambda(n'')$ and thus $\lambda(n''+1) = f(\lambda(n')) = f(n, 0) = (k, m)$ so we have found a $n' = n''+1$ such that $\lambda(n') = (k, m)$

   b. **($k \neq 0$)** then $0 < k \Rightarrow k-1$ exists and as $0 \leqslant m \Rightarrow 0 < m+1$ we have $f(k-1, m+1) = ((k-1)+1, (m+1)-1) = (k, m)$. Now $((k+m)-1)+0 = (n+1) - 1 = n \in S$ so that $\exists n'' \in \mathbb{N}_0$ such that $\lambda(n'') = ((k+m)-1, 0)$. So $\lambda(n''+1) = f((k+m) - 1, 0) = (0, k+m)$. Using **property a** of λ we have then that $\lambda((n''+1)+k) = (k, m)$ so we have found a $n' = (n''+1)+k$ such that $\lambda(n') = (k, m)$

If thus $(k, m) \in \mathbb{N}_0 \times \mathbb{N}_0$ then $n = k+m \in \mathbb{N}_0 = S \Rightarrow \exists n' \in \mathbb{N}_0 \vdash \lambda(n') = (k, m)$ proving that $\lambda: \mathbb{N}_0 \times \mathbb{N}_0 \rightarrow \mathbb{N}_0$ is surjective and as we have by construction of λ it is injective we have that $\lambda: \mathbb{N}_0 \times \mathbb{N}_0 \rightarrow \mathbb{N}_0$ is a bijection and thus $\mathbb{N}_0 \approx \mathbb{N}_0 \times \mathbb{N}_0$ □

Corollary 5.59. *If A, B are **denumerable** then $A \times B$ is denumerable*

Proof. If A, B are denumerable then $\mathbb{N}_0 \approx A$ and $\mathbb{N}_0 \approx B$ we have by 5.12 that $A \times B \approx \mathbb{N}_0 \times \mathbb{N}_0 \approx \mathbb{N}_0$ so $A \times B \approx \mathbb{N}_0$ □

Theorem 5.60. *If $\{A_i\}_{i \in B}$ is a **denumerable** family of **denumerable** sets (so B is denumerable and $\forall i \in B$ we have that A_i is a denumerable set) then $\bigcup_{i \in B} A_i$ is **denumerable**.*

Proof. As $\{A_i\}_{i \in B}$ is a denumerable family of denumerable sets there exists a graph A with $\mathrm{dom}(G) = B$ where $\mathbb{N}_0 \approx B$, so there is a bijection $f: \mathbb{N}_0 \rightarrow B$, and $\forall i \in B$ we have $A_i \approx \mathbb{N}_0 \Rightarrow \forall i \in B$ there exists a bijection $f_i: \mathbb{N}_0 \rightarrow A_i$. So define $\forall i \in B$ the non empty set (see 2.72) $\mathcal{A}_i = \{f \in A_i^{\mathbb{N}_0}: f \text{ is } a \text{ bijection}\}$, then $\mathcal{A} = \bigcup_{i \in B} \mathcal{A}_i$ is a set and $\forall i \in B$ we have $\mathcal{A}_i \subseteq \mathcal{A}$. Using the axiom of choice (see 2.198) we have then the existence of a choice function $c: \mathcal{P}'(\mathcal{A}) \rightarrow \mathcal{A}$ with $c(\mathcal{A}_i) \in \mathcal{A}_i \Rightarrow c(\mathcal{A}_i)$ is a bijection from $\mathbb{N}_0 \rightarrow A_i$. Define now $F: \mathbb{N}_0 \times \mathbb{N}_0 \rightarrow \bigcup_{i \in B} A_i$ by $F(n, m) = c(\mathcal{A}_{f(n)})(m)$ then F is surjective. For if $y \in \bigcup_{i \in B} A_i$ then there exists a $i \in B$ such that $y \in A_i$ and as f is bijective there exists a $n \in \mathbb{N}_0$ such that $f(n) = i$, also as $c(A_{f(n)})$: $\mathbb{N}_0 \rightarrow A_{f(n)} = A_i$ there exists a $m \in \mathbb{N}_0$ such that $c(A_{f(n)})(m) = y$ and thus $F(n, m) = c(\mathcal{A}_{f(n)})(m)$. As $\mathrm{N} \times \mathbb{N}_0 \approx \mathbb{N}_0$ there exists a bijection $\varphi: \mathbb{N}_0 \rightarrow \mathbb{N}_0 \times \mathbb{N}_0$ and thus $F \circ \varphi: \mathbb{N}_0 \rightarrow \bigcup_{i \in B} A_i$ is a surjective function. Using 5.9 we have then that $\bigcup_{i \in B} A_i \approx E \subseteq \mathbb{N}_0$. Using 5.55 we have that E is finite or denumerable hence $\bigcup_{i \in B} A_i$ is either finite or denumerable. As $A_{f(1)} \subseteq \bigcup_{i \in B} A_i$ and $A_{f(1)}$ is denumerable $\underset{5.34}{\Rightarrow} A_{f(1)}$ is infinite $\underset{5.35}{\Rightarrow} \bigcup_{i \in B} A_i$ is not finite so $\bigcup_{i \in B} A_i$ is a denumerable set. □

Corollary 5.61. *The union of two **denumerable** sets is **denumerable**.*

Proof. If A, B are denumerable sets then there exists bijections $f: \mathbb{N}_0 \rightarrow A$ and $g: \mathbb{N}_0 \rightarrow B$. Define now $F: \{1,2\} \times \mathbb{N}_0 \rightarrow A \bigcup B$ by $F(1,n) = f(n)$ and $F(2,n) = g(n)$ then F is a surjection as if $x \in A \bigcup B$ we have either $x \in A \underset{f \text{ is bijective}}{\Rightarrow} \exists n \in \mathbb{N}_0 \vdash x = f(n) = F(1,n)$ or $x \in B \underset{g \text{ is bijective}}{\Rightarrow} \exists n \in \mathbb{N}_0 \vdash x = g(n) F(2,n)$. Now $\{1.2\} \times \mathbb{N}_0 \subseteq \mathbb{N}_0 \times \mathbb{N}_0$ then by 5.55 and the fact that $\mathbb{N}_0 \approx \mathbb{N}_0 \times \mathbb{N}_0$ we have that $\{1,2\} \times \mathbb{N}_0$ is either finite or denumerable, as $\mathbb{N}_0 \approx \{1\} \times \mathbb{N}_0 \subseteq \{1,2\} \times \mathbb{N}_0$ we have that $\{1,2\} \times \mathbb{N}_0$ is infinite and thus is denumerable. So there exists a bijection $\varphi: \mathbb{N}_0 \rightarrow \{1,2\} \times \mathbb{N}_0$ and thus a surjection $F \circ \varphi: \mathbb{N}_0 \rightarrow A \bigcup B$ which means by 5.9 that $A \bigcup B \approx E \subseteq \mathbb{N}_0$. Using 5.55 we have that E is finite or denumerable hence $A \bigcup B$ is either finite or denumerable, as $A \subseteq A \bigcup B$ is a infinite subset we must have that $A \bigcup B$ is not finite and thus that $A \bigcup B$ is denumerable. $\square$

Using mathematical induction we can extend this to a finite union of denumerable sets

Corollary 5.62. *Let $n \in \mathbb{N}$, $\{A_i\}_{i \in \{1,\ldots,n\}}$ a finite family of denumerable sets then $\bigcup_{i \in \{1,\ldots,n\}} A_i$ is denumerable.*

Proof. We prove this by induction so let $\mathcal{S} = \left\{n \in \mathbb{N} | \forall \{A_i\}_{i \in \{1,\ldots,n\}} | \forall i \in \{1,\ldots,n\}\ A_i \text{ is denumerable we have that } \bigcup_{i \in \{1,\ldots,n\}} A_i \text{ is denumerable}\right\}$ then we have

- **$1 \in \mathcal{S}$.** If $\{A_i\}_{i \in \{1,\ldots,1\}}$ is a family of denumerable sets then $\bigcup_{i \in \{1,\ldots,1\}} A_i = A_1$ is denumerable proving that $1 \in \mathcal{S}$
- **$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$.** Let $\{A_i\}_{i \in \{1,\ldots,n+1\}}$ be a family of denumerable sets then as $n \in \mathcal{S}$ we have that $\bigcup_{i \in \{1,\ldots,n\}} A_i$ is denumerable. So $\bigcup_{i \in \{1,\ldots,n+1\}} A_i = \left(\bigcup_{i \in \{1,\ldots,n\}} A_i\right) \bigcup A_{n+1}$ is denumerable by 5.61, proving that $n+1 \in \mathcal{S}$. $\square$

As also the finite union of finite sets is finite (see 5.39) we have the following corollary of the above corollary.

Corollary 5.63. *The finite union of* ***countable*** *sets is* ***countable****.*

Corollary 5.64. *Let A be a non empty finite set, B a denumerable set then $A \times B$ is denumerable*

Proof. First as A is finite and non empty there exists a bijection $\sigma: \{1,\ldots,n\} \rightarrow A$ where $n \in \mathbb{N}$. So we have if $x = (x_1, x_2) \in A \times B$ then $x_1 \in A \wedge x_2 \in B \underset{\exists i \in \{1,\ldots,n\} \vdash \sigma(i) = x_1}{\Rightarrow} x_1 = \sigma(i) \in \{\sigma(i)\} \wedge x_2 \in B$, proving that $x = (x_1, x_2) \in \bigcup_{i \in \{1,\ldots,n\}} (\{\sigma(i)\} \times B)$. If $x = (x_1, x_2) \in \bigcup_{i \in \{1,\ldots,n\}} (\{\sigma(i)\} \times B) \Rightarrow \exists i \in \{1,\ldots,n\}$ so that $x_1 = \sigma(i) \subseteq A \wedge x_2 \in B \Rightarrow x = (x_1, x_2) \in A \times B$. So we have

$$\bigcup_{i \in \{1,\ldots,n\}} (\{\sigma(i)\} \times B) = A \times B \tag{5.1}$$

As B is denumerable there exists a bijection $\tau: \mathbb{N} \rightarrow B$. Then $\forall i \in \{1,\ldots,n\}$ we have that $\rho_i: \mathbb{N} \rightarrow \{\sigma(i)\} \times B$ defined by $\tau(j) = (\sigma(i), \tau(i))$ is bijection for we have

- **injectivity.** If $\rho_i(k) = \rho_i(l)$ then $(\sigma(i), \tau(k)) = (\sigma(i), \tau(l)) \Rightarrow \tau(k) = \tau(l) \underset{\tau \text{ is bijection}}{\Rightarrow} k = l$

surjectivity. If $(x,y)\in\{\sigma(i)\}\times B$ then as τ is a bijection there exists a $k\in\mathbb{N}$ such that $\tau(k)\in B$ so that $\rho_i(k)=(\sigma(i),\tau(k))\in\{\sigma(i)\}\times B$.

The above proves that $\forall i\in\{1,\ldots,n\}$ we have that $\{\sigma(i)\}\times B$ is denumerable, the theorem follows then from 5.62 and 5.1. □

5.4 Some properties of countable sets

Using the definition of **countability** and 5.55 we have the following theorem

Theorem 5.65. *Every subset of a **denumerable** set is **countable**.*

As the subsets of finite sets are finite we have that the following is true.

Corollary 5.66. *Every subset of a **countable** set is **countable**.*

Theorem 5.67. *If X is a non-empty set then the following are equivalent*

1. *X is countable*
2. *There exists a surjective function $f:\mathbb{N}_0\rightarrow X$*
3. *There exists a injective function $f:X\rightarrow\mathbb{N}_0$*

Proof.

1. **($\mathbf{1\Rightarrow 2}$)** If X is countable then we have the following possibilities
   a. **($\boldsymbol{X}$ is finite)** Then $\exists n\in\mathbb{N}_0$ such that $n\approx X$ so there exists a bijection $f':n=S_n\rightarrow X$, also as $X\neq\emptyset$ there exists a $x\in X$. Define now f: $\mathbb{N}_0\rightarrow X$ by

   $$f(n) = \begin{cases} f'(n)\text{ if } i<n\text{ (or } i\in S_n) \\ x\text{ if } n\leqslant i \end{cases}$$

   which is surjective because f' is surjective.

   b. **($\boldsymbol{X}$ is infinite)** Then $\mathbb{N}_0\approx X\Rightarrow$ there exists a bijection (thus surjection) between $\mathbb{N}_0$ and X
2. **($\mathbf{2\Rightarrow 3}$)** If a surjective function $f:\mathbb{N}_0\rightarrow X$ exists then by 2.199 there exists a injective function $g:X\rightarrow\mathbb{N}_0$
3. **($\mathbf{3\Rightarrow 1}$)** If X is finite then it is of course countable. So assume that X is infinite. Then from the existence of the injective function $f:X\rightarrow\mathbb{N}_0$ we have that $f:X\rightarrow f(X)$ is a bijection and thus $f(X)$ is infinite. From 5.54 we have then that $f(X)$ is denumerable$\Rightarrow X\approx f(X)\approx\mathbb{N}_0\Rightarrow X\approx\mathbb{N}_0$ and thus X is countable. □

Lemma 5.68. *If $n\in\mathbb{N}_0$ then $n\times\mathbb{N}_0$ and $\mathbb{N}_0\times n$ are **countable**. If $n\neq 0$ then $n\times\mathbb{N}_0$ and $\mathbb{N}_0\times n$ are denumerable.*

Proof. First $f: n \times \mathbb{N}_0 \rightarrow \mathbb{N}_0 \times n$ define by $(i, j) \Rightarrow f(i, j) = (j, i)$ is bijective

1. **(injectivity)** If $f(i, j) = f(i', j') \Rightarrow (j, i) = (j', i') \Rightarrow i = i' \wedge j = j' \Rightarrow (i, j) = (i', j')$
2. **(surjectivity)** If $(i, j) \in \mathbb{N}_0 \times n \Rightarrow i \in \mathbb{N}_0 \wedge j \in n \Rightarrow (j, i) \in n \times N$ and $f(j, i) = (i, j)$

so we have $n \times \mathbb{N}_0 \approx \mathbb{N}_0 \times n$. We must thus only prove that $n \times \mathbb{N}_0$ is denumerable. First $n \times \mathbb{N}_0 = S_n \times \mathbb{N}_0 \subseteq \mathbb{N}_0 \times \mathbb{N}_0$ so we have by 5.55 that $n \times \mathbb{N}_0$ is either finite or denumerable and is thus countable. If $n \neq 0$ then $0 < n = S_n \Rightarrow 0 \in n \Rightarrow \{0\} \subseteq \mathbb{N}_0 \Rightarrow \{0\} \times \mathbb{N}_0 \subseteq n \times \mathbb{N}_0$. And as it is trivial to prove that $f: \mathbb{N}_0 \rightarrow \{0\} \times \mathbb{N}_0: i \rightarrow (0, i)$ is a bijection we have that $\{0\} \times \mathbb{N}_0$ is denumerable and thus by 5.34 it is infinite, so by 5.35 $n \times \mathbb{N}_0$ is not finite and must thus be denumerable. $\square$

Theorem 5.69. *If A, B are countable sets then $A \times B$ is countable*

Proof. If A, B are countable sets then we have the following cases to consider

1. **($\boldsymbol{A}$ finite $\wedge$ $\boldsymbol{B}$ finite)** then $A \times B$ is finite by 5.44 and thus countable.
2. **($\boldsymbol{A}$ finte $\wedge$ $\boldsymbol{B}$ denumerable)** Then $\exists n \in \mathbb{N}_0$ such that $A \approx n$ and $B \approx \mathbb{N}_0 \Rightarrow A \times B \approx n \times \mathbb{N}_0$ which by the above lemma is countable so we have that $A \times B$ is countable.
3. **($\boldsymbol{A}$ denumerable $\wedge$ $\boldsymbol{B}$ finite)** Then $\exists n \in \mathbb{N}_0$ such that $A \approx \mathbb{N}_0$ and $B \approx n \Rightarrow A \times B \approx \mathbb{N}_0 \times n$ which by the above lemma is countable;e so we have that $A \times B$ is countable.
4. **($\boldsymbol{A}$ denumerable $\wedge$ $\boldsymbol{B}$ denumerable)** Then by 5.59 we have that $A \times B$ is denumerable and thus countable. $\square$

Lemma 5.70. *If $\{A_i\}_{i \in S_n}$ is a family of countable sets then $\bigcup_{i \in S_n} A_i$ is countable.*

Proof. We prove this by induction on n. So define $S = \{n \in \mathbb{N}_0 | \text{if } \{A_i\}_{i \in S_n} \text{ is a family of countable sets then } \bigcup_{i \in S_n} A_i \text{ is countable}\}$ then we have

1. If $n = 0$ take then $\{A_i\}_{i \in S_0}$ then there exists a graph A with $\mathrm{dom}(A) = \emptyset$ so if $x \in \bigcup_{i \in S_0} A_i$ then there exists a $i \in S_0$ such that $x \in A_i$ which is a contradiction as $S_0 = \emptyset$ so we have that $\bigcup_{i \in S_0} A_i = \emptyset$ and thus finite and countable. So we have $0 \in S$
2. Assume that $n \in S$ and take $n+1$ then we have for the family of countable sets $\{A_i\}_{i \in S_{n+1}}$ that there exists a graph A with $\mathrm{dom}(A) = S_{n+1}$. If we take $A' = \{(x, y) \in A | x \in S_n\}$ (so $\mathrm{dom}(A') = S_n$) then $\bigcup_{i \in S_{n+1}} A_i = A_n \bigcup \left(\bigcup_{i \in S_n} A'_i\right)$

   Proof.

   a. If $x \in \bigcup_{i \in S_{n+1}} A_i$ then $\exists i \in \mathrm{dom}(A) = S_{n+1} \Rightarrow i < n+1$ such that $x \in A_i$ then we have the following cases

      i. **($\boldsymbol{i = n}$)** then $x \in A_n \Rightarrow x \in A_n \bigcup \left(\bigcup_{i \in S_n} A'_i\right)$

ii. $(\boldsymbol{i \neq n})$ then $i<n$ [if $n \leqslant i \underset{i \neq n}{\Rightarrow} n<i<n+1 \Rightarrow n+1 \leqslant i<n+1 \Rightarrow n+1 \neq n+1$ a contradiction] and thus $i \in S_n$. As $x \in A_i$ we have $(i,x) \in A \underset{i \in S_n}{\Rightarrow} (i,x) \in A' \Rightarrow x \in \bigcup_{i \in S_n} A'_i \Rightarrow x \in A_n \bigcup (\bigcup_{i \in S_n} A'_i)$

in all cases we have $x \in A_n \bigcup (\bigcup_{i \in S_n} A'_i)$

b. If $x \in A_n \bigcup (\bigcup_{i \in S_n} A'_i)$ then we have the following cases

i. $(\boldsymbol{x \in A_n}) \Rightarrow x \in \bigcup_{i \in S_{n+1}} A_i$

ii. $(\boldsymbol{x \notin A_n}) \Rightarrow x \in \bigcup_{i \in S_n} A'_i \Rightarrow \exists i \in S_n \vdash x \in A'_i \Rightarrow (i,x) \in A' \subseteq A \Rightarrow x \in A_i \Rightarrow x \in \bigcup_{i \in S_{n+1}} A_i$

in all cases $x \in \bigcup_{i \in S_{n+1}} A_i$ □

As $n \in S$ we have $\bigcup_{i \in S_n} A'_i$ is countable and as by assumption A_n is finite we have by 5.63 that $\bigcup_{i \in S_{n+1}} A_i = A_n \bigcup (\bigcup_{i \in S_n} A'_i)$ is finite and thus $n+1 \in S$.

By induction we have then $S = \mathbb{N}_0$ proving our theorem. □

Lemma 5.71. *If $\{A_i\}_{i \in I}$ is a **finite** family (I is finite) of **countable** sets then $\bigcup_{i \in I} A_i$ is **countable**.*

Proof. As I is finite we have the existence of a bijection $f: n = S_n \rightarrow I$ $(n \in \mathbb{N}_0)$ and thus using 2.64 we have that $\bigcup_{i \in I} A_i = \bigcup_{j \in S_n} A_{f(j)}$ which is countable by the previous lemma. □

Theorem 5.72. *If $\{A_i\}_{i \in B}$ is a **denumerable** family of **countable** sets (so B is denumerable and $\forall i \in B$ we have that A_i is a **countable** set) then $\bigcup_{i \in B} A_i$ is **countable**. Further if we have in addition that $\forall i \in B \quad A_i \neq \emptyset$ and $\forall i,j \in B$ with $i \neq j$ that $A_i \bigcap A_j = \emptyset$ then $\bigcup_{i \in B} A_i$ is denumerable.*

Proof. As $\{A_i\}_{i \in B}$ is a denumerable family of countable sets, there exists a graph A with $\operatorname{dom}(G) = B$ where $\mathbb{N}_0 \approx B$, so there is a bijection $f: \mathbb{N}_0 \rightarrow B$, and $\forall i \in B$ we have that A_i is countable. So $\forall i \in B$ there exists a bijection $g: N \rightarrow A_i$ where N is either $\mathbb{N}_0$ or $n \in \mathbb{N}_0$. So if we define $\forall i \in B$ the non empty set (see 2.72) $\mathcal{A}_i = \{f \in A_i^{\mathbb{N}_0} | f \text{ is a bijection}\} \bigcup (\bigcup_{n \in \mathbb{N}_0} \{f \in A_i^n | f \text{ is a bijection}\})$, then $\mathcal{A} = \bigcup_{i \in B} \mathcal{A}_i$ is a set and $\forall i \in B$ we have $\mathcal{A}_i \subseteq \mathcal{A}$. Using the axiom of choice (see 2.198) we have then the existence of a choice function $c: \mathcal{P}'(\mathcal{A}) \rightarrow \mathcal{A}$ so that $c(\mathcal{A}_i) \in \mathcal{A}_i \Rightarrow c(\mathcal{A}_i)$ is a bijection from $\mathbb{N}_0 \rightarrow A_i$ or a bijection from $n \rightarrow A_i$. Define now F: $\mathbb{N}_0 \times \mathbb{N}_0 \rightarrow \bigcup_{i \in B} A_i$ by $F(n,m) = c(\mathcal{A}_{f(n)})(m)$.

We can then prove that F is surjective. For if $y \in \bigcup_{i \in B} A_i \Rightarrow \exists i \in B \vdash y \in A_i \underset{f: \mathbb{N}_0 \rightarrow B \text{ is bijective}}{\Rightarrow} \exists n \in \mathbb{N}_0 \vdash f(n) = i \Rightarrow y \in A_{f(n)}$. We have now the following cases for $c(A_i) = c(A_{f(n)})$:

1. $(\boldsymbol{\exists m \in \mathbb{N}_0 \vdash c(A_{f(n)}): m \rightarrow A_{f(n)}}$ **is a bijection**$) \Rightarrow \exists j \in m \vdash c(A_{f(n)})(j) = y \Rightarrow F(n,j) = c(A_{f(n)})(j) = y$

2. **($c(A_{f(n)})$: $\mathbb{N}_0 \rightarrow A_{f(n)}$ is a bijection)** $\Rightarrow \exists j \in \mathbb{N}_0 \vdash c(A_{f(n)})(j) = y \Rightarrow F(n, j) = c(A_{f(n)})(j) = y$

proving that F is surjective.

As $\mathbb{N} \times \mathbb{N}_0 \approx \mathbb{N}_0$ there exists a bijection $\varphi: \mathbb{N}_0 \rightarrow \mathbb{N}_0 \times \mathbb{N}_0$ and thus $F \circ \varphi: \mathbb{N}_0 \rightarrow \bigcup_{i \in B} A_i$ is a surjective function. Using 5.9 we have then that $\bigcup_{i \in B} A_i \approx E \subseteq \mathbb{N}_0$. Using 5.55 we have that E is finite or denumerable hence $\bigcup_{i \in B} A_i$ is either finite or denumerable and thus countable.

Finally if in addition $\forall i \in B$ $A_i = \emptyset$ and $\forall i, j \in B \vdash i \neq j$ we have $A_i \bigcap A_j$. Using the axiom of choice (see 2.198) there exists a function $\tau: B \rightarrow \bigcup_{i \in B} A_i$ such that $\forall i \in B$ we have $\tau(i) \in A_i$. If $\tau(i) = \tau(j)$ then $\tau(i) \in A_i \wedge \tau(i) = \tau(j) \in A_j \Rightarrow \tau(i) \in A_i \bigcap A_j$ from which it follows that $i = j$ [if $i \neq j$ then $\tau(i) \in A_i \bigcap A_j = \emptyset$ a contradiction), hence $\tau: B \rightarrow \bigcup_{i \in B} A_i$ is a injection. From this it follows that $\tau: B \rightarrow \tau(B) \subseteq \bigcup_{i \in B} A_i$ is a bijection, hence $\tau \circ f: \mathbb{N}_0 \rightarrow \tau(B)$ is a bijection which means that $\tau(B)$ is denumerable and as $\tau(B) \subseteq \bigcup_{i \in B} A_i$ we have by 5.34 that $\bigcup_{i \in B} A_i$ is infinite and as it is countable it must be finite. □

Using the previous lemma and the above theorem we have then the following theorem.

Theorem 5.73. *If $\{A_i\}_{i \in B}$ is a* ***countable*** *family of* ***countable*** *sets then we have that $\bigcup_{i \in B} A_i$ is countable.*

Proof. If $\{A_i\}_{i \in B}$ is a countable family of countable sets then we have either:

1. **(B is finite)** then by 5.71 we have that $\bigcup_{i \in B} A_i$ is countable.
2. **(B is denumerable)** then by 5.73 we have that $\bigcup_{i \in B} A_i$ is countable. □

5.5 Sequences

Definition 5.74. *If X is a set then a family $\{x_i\}_{i \in \mathbb{N}}$ of elements in X is called a sequence.*

Proposition 5.75. *Let $\langle X, \leqslant \rangle$ be a paritally ordered set and $\{x_i\}_{i \in \mathbb{N}} \subseteq X$ a sequence of elements in X such that $\forall i \in \mathbb{N}$ we have $x_i \leqslant x_{i+1}$ then $\forall i, j \in \mathbb{N}$ with $i < j$ we have $x_i \leqslant x_j$.*

Proof. We prove this by induction so given $i \in \mathbb{N}$ take $\mathcal{S}_i = \{n \in \mathbb{N} | x_i \leqslant x_{i+n}\}$ then we have

$1 \in \mathcal{S}_i$. by assumption we have $x_i \leqslant x_{i+1}$ proving that $1 \in \mathcal{S}_i$

$n \in \mathcal{S}_i \Rightarrow n + 1 \in \mathcal{S}_i$. as $n \in \mathcal{S}_i$ we have $x_i \leqslant x_{i+n}$ and by the hypothesis we have $x_{i+n} \leqslant x_{(i+n)+1} \leqslant x_i + (n+1)$ so by transitivity we have $x_i \leqslant x_{i+(n+1)}$ proving that $n + 1 \in \mathcal{S}_i$

Mathematical induction proves that $\mathbb{N} = \mathcal{S}_i$. So if $i, j \in \mathbb{N}$ with $i < j$ then $j - i \in \mathbb{N} = \mathcal{S}_i$ and $x_i \leqslant x_{i+(j-i)} = x_j$ proving the proposition. □

Proposition 5.76. *Let $\langle X, \leqslant \rangle$ be a paritally ordered set and $\{x_i\}_{i \in \mathbb{N}} \subseteq X$ a sequence of elements in X such that $\forall i \in \mathbb{N}$ we have $x_{i+1} \leqslant x_i$ then $\forall i, j \in \mathbb{N}$ with $i < j$ we have $x_j \leqslant x_i$.*

Proof. We prove this by induction so given $i \in \mathbb{N}$ take $\mathcal{S}_i = \{n \in \mathbb{N} | x_{i+n} \leqslant x_i\}$ then we have

$1 \in \mathcal{S}_i$. by assumption we have $x_{i+1} \leqslant x_i$ proving that $1 \in \mathcal{S}_i$

$n \in \mathcal{S}_i \Rightarrow n + 1 \in \mathcal{S}_i$. as $n \in \mathcal{S}_i$ we have $x_{i+n} \leqslant x_i$ and by the hypothesis we have $x_{i+(n+1)} = x_{(i+n)+1} \leqslant x_i$ so by transitivity we have $x_{i+(n+1)} \leqslant x_i$ proving that $n + 1 \in \mathcal{S}_i$

Mathematical induction proves that $\mathbb{N} = \mathcal{S}_i$. So if $i, j \in \mathbb{N}$ with $i < j$ then $j - i \in \mathbb{N} = \mathcal{S}_i$ and $x_j = x_{i+(j-i)} \leqslant x_i$ proving the proposition. □

Corollary 5.77. *Let $\mathcal{A}$ be a class of classes and $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$ such that $\forall i \in \mathbb{N}$ we have $A_i \subseteq A_{i+1}$ [or $A_{i+1} \subseteq A_i$] then $\forall i, j \in \mathbb{N}$ with $i < j$ we have $A_i \subseteq A_j$ [or $A_j \subseteq A_i$]*

Proof. This follows because $\langle \mathcal{A}, \subseteq \rangle$ is a partially ordered class (see 2.133) so that we can apply 5.75 and 5.76. □

5.6 Finite Cartesian product of sets

First some notations about the generalized product of sets in the finite case (see 2.77 and 2.84)

Notation 5.78. *Let $n, m \in \mathbb{N}$, $\{A_i\}_{i \in \{n, \ldots, m\}}$ be a finite family of sets then we note $x \in \prod_{i \in \{n, \ldots, m\}} A_i$ by $x = (x_n, \ldots, x_m)$ meaning $x = \{(n, x_n), \ldots, (m, x_m)\} \subseteq \{n, \ldots, m\} \times \left(\bigcup_{i \in \{n, \ldots, m\}} A_i\right)$ such that $\langle x, \{n, \ldots, m\} \rangle$ is a pretuple and $\forall i \in \{n, \ldots, m\}$ we have $x_i \in A_i$ (see the definition of $\prod_{i \in \{n, \ldots, m\}} A_i$ at 2.77)*

Using the above notation we can rephrase theorem 2.99 as follows

Theorem 5.79. *Let $n, k \in \mathbb{N}$ such that $1 \leqslant k < n$ and $\{A_i\}_{i \in \{n, \ldots, m\}}$ a finite family then given $x \in \prod_{i \in \{1, \ldots, n\}} A_i$ we have that $x_{|\{1, \ldots, k\}} \in \prod_{i \in \{1, \ldots k\}} A_i$, $x_{|\{k+1, \ldots, n\}} \in \prod_{i \in \{k+1, \ldots, n\}} A_i$ and $\beta: \prod_{i \in \{1, \ldots, n\}} A_i \rightarrow \left(\prod_{i \in \{1, \ldots, k\}} A_i\right) \times \left(\prod_{i \in \{k+1, \ldots, n\}} A_i\right)$ defined by $\beta(x) = (x_{|\{1, \ldots, k\}}, x_{|(k+1, \ldots, n)})$ is a bijection. Or using the above notation we reformulate that as, if $(x_1, \ldots, x_n) \in \prod_{i \in \{1, \ldots, n\}} A_i$ then $(x_1, \ldots, x_k) \in \prod_{i \in \{1, \ldots, k\}} A_i$, $(x_{k+1}, \ldots, x_n) \in \prod_{i \in \{k+1, \ldots, n\}} A_i$ and $\beta((x_1, \ldots, x_n)) = ((x_1, \ldots, x_k), (x_{k+1}, \ldots, x_n))$.*

The next theorem shows how to extend a tuple

Theorem 5.80. *Let $n\in\mathbb{N}$, $\{A_i\}_{i\in\{1,\ldots,n+1\}}$ be a finite family of sets and consider $\{A_i\}_{i\in\{1,\ldots,n\}}$ the finite subfamily formed by restriction to $\{1,\ldots,n\}$ then if $x\in\prod_{i\in\{1,\ldots,n\}}A_i$, $y\in A_{n+1}$ define $x_{+y}=x\bigcup\{(n+1,y)\}$ then $x_{+y}\in\prod_{i\in\{1,\ldots,n+1\}}A_i$ and $(x_{+y})_{|\{1,\ldots,n\}}=x$.*

Proof. As $x\subseteq\{1,\ldots,n\}\times\left(\bigcup_{i\in\{1,\ldots,n\}}A_i\right)$ we have then if $(i,z)\in x_+$ either

- **$(i,z)\in x$.** so that $(i,z)\in\{1,\ldots,n\}\times\left(\bigcup_{i\in\{1,\ldots,n\}}A_i\right)\subseteq_{1.44}\{1,\ldots,n+1\}\times\left(\bigcup_{i\in\{1,\ldots,n\}}A_{n+1}\right)$
- **$(i,z)\in\{(n+1,y)\}$.** so that $i=n+1\wedge z=y\underset{y\in A_{n+1}}{\Rightarrow}i\in\{1,\ldots,n+1\}\wedge y\in\bigcup_{i\in\{1,\ldots,n+1\}}A_i\Rightarrow(i,z)\in\{1,\ldots,n+1\}\times\left(\bigcup_{i\in\{1,\ldots,n+1\}}A_i\right)$

so we have in all cases $(i,y)\in\{1,\ldots,n+1\}\times\left(\bigcup_{i\in\{1,\ldots,n+1\}}A_i\right)$ proving that

$$x_{+y}\subseteq\{1,\ldots,n+1\}\times\left(\bigcup_{i\in\{1,\ldots,n+1\}}A_i\right). \tag{5.2}$$

Also if $(i,z),(i,z')\in x_{+y}$ we have the following cases to consider then

- **$i\in\{1,\ldots,n\}$.** then $(i,z),(i,z')\in x\underset{\langle x,\{1,\ldots,n\}\rangle\text{ is a pretuple}}{\Rightarrow}z=z'$
- **$i=n+1$.** then $z=y=z'$

proving that x_{+y} is a function graph. If $i\in\{1,\ldots,n+1\}$ then if $i=n+1$ we have $(n+1,y)\in x_{+y}$ and if $i\in\{1,\ldots,n\}$ there exists as $\operatorname{dom}(x)=\{1,\ldots,n\}$ a z such that $(i,z)\in x\subseteq x_{+y}$ proving that $\operatorname{dom}(x_{+y})=\{1,\ldots,n+1\}$. So we have proved that

$$\langle x_{+y},\{1,\ldots,n+1\}\rangle\text{ is a pretuple} \tag{5.3}$$

If $i\in\{1,\ldots,n+1\}$ then either $i\in\{1,\ldots,n\}\Rightarrow(i,x_{+y}(i))\in x\Rightarrow x_{+y}(i)=x(i)\in A_i$ or $i=n+1$ then $x_{+y}(i)=x_{+y}(n+1)=y\in A_{n+1}$ so that

$$\forall i\in\{1,\ldots,n+1\}\text{ we have }x_{+y}(i)\in A_i \tag{5.4}$$

Using 5.2,5.3 and 5.4 we have that $x_{+y}\in\prod_{i\in\{1,\ldots,n+1\}}A_i$ as also $(x_{+y})_{|[1,\ldots,n]}=\{(i,z)\in x_{+y}|i\in\{1,\ldots,n\}\}=x$ we have proved the theorem. □

Next we define the finite Cartesian product of sets and show its relation with the normal product of a family of sets (see 2.77).

Definition 5.81. (finite Cartesian product of sets) *Let $n\in\mathbb{N}$ and $\{A_i\}_{i\in\{1,\ldots,n\}}$ a finite family of sets then we define $\bigotimes_{i\in\{1,\ldots,n\}}A_i$ recursively as follows:*

- **$n=1$.** $\bigotimes_{i\in\{1,\ldots,1\}}A_i=A_1$
- **$n>1$.** $\bigotimes_{i\in\{1,\ldots,n\}}A_i=\left(\bigotimes_{i\in\{1,\ldots,n-1\}}A_i\right)\times A_n$ *(see 1.41)*

Example 5.82. Let $\{A_i\}_{i\in\{1,\ldots,3\}}$ be a finite family of sets then $\bigotimes_{i\in\{1,\ldots,3\}}A_i=\left(\bigotimes_{i\in\{1,\ldots,2\}}A_i\right)\times A_3=\left(\left(\bigotimes_{i\in\{1\ldots,1\}}A_i\right)\times A_2\right)\times A_3=(A_1\times A_2)\times A_3$ and $x\in\bigotimes_{i\in\{1,\ldots3\}}A_i\Leftrightarrow\exists x_1\in A_1,x_2\in A_2,x_3\in A_3$ so that $x=((x_1,x_2),x_3)$

Theorem 5.83. *Let* $\{A_i\}_{i\in\{1,\ldots,n\}}, \{B_i\}_{i\in\{1,\ldots,n\}}$ *be finite families of sets such that* $\forall i\in\{1,\ldots,n\}$ *we have* $B_i\subseteq A_i$ *then* $\bigotimes_{i\in\{1,\ldots,n\}} B_i\subseteq\bigotimes_{i\in\{1,\ldots,n\}} A_i$

Proof. We prove this by induction, so let $\mathcal{S}=\left\{n\in\mathbb{N}|\text{If}\{B_i\}_{i\in\{1,\ldots,n\}}, \{A_i\}_{i\in\{1,\ldots,n\}}\right.$ are finite families of sets such that $\forall i\in\{1,\ldots,n\}\ B_i\subseteq A_i$ then $\bigotimes_{i\in\{1,\ldots,n\}} B_i\subseteq \left.\bigotimes_{i\in\{1,\ldots,n\}} A_i\right\}$. Then we have

$1\in\mathcal{S}$. If $x\in\bigotimes_{i\in\{1,\ldots,1\}} B_i=B_1\subseteq A_1=\bigotimes_{i\in\{1,\ldots,1\}} A_i$ proving that $\bigotimes_{i\in\{1,\ldots,1\}} B_i\subseteq\bigotimes_{i\in\{1,\ldots,1\}} A_i$, hence $1\in\mathcal{S}$

$n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$. Let $x\in\prod_{i\in\{1,\ldots,n+1\}} B_i$ then as $n+1>1$ then $\exists y\in\bigotimes_{i\in\{1,\ldots,n\}} B$ and $\exists z\in B_{n+1}$ such that $x=(y,z)$, As $n\in\mathcal{S}$ we have that $y\in\bigotimes_{i\in\{1,\ldots,n\}} A_i$, further $z\in B_{n+1}\subseteq A_{n+1}$. Applying 1.44 gives $x=(y,z)\in\left(\bigotimes_{i\in\{1,\ldots,n\}} A_i\right)\times A_{n+1}=\bigotimes_{i\in\{1,\ldots,n+1\}} A_i$ proving that $\bigotimes_{i\in\{1,\ldots,n+1\}} B_i\subseteq\bigotimes_{i\in\{1,\ldots,n+1\}} A_i$. Hence $n+1\in\mathcal{S}$. □

Next we show that with every $x\in\prod_{i\in\{1,\ldots,n\}} A_i$ we can construct a $[x]\in\bigotimes_{i\in\{1,\ldots,n\}} A_i$

Definition 5.84. (product tuple) *Let* $n\in\mathbb{N}$, $A=\{A_i\}_{i\in\{1,\ldots,n\}}$ *be a finite family of sets,* $x\in\prod_{i\in\{1,\ldots,n\}} A_i$ *then we define* $[x]_n\in\bigotimes_{i\in\{1,\ldots,n\}} A_i$ *recursively by*

$n=1$. $[x]_1=\pi_1(x)$

$n>1$. $[x]_n=([x_{|\{1,\ldots,n-1\}}]_{n-1},\pi_n(x))$

where π_i *is the projection function defined in 2.87.*

Example 5.85. For $1\in\mathbb{N}$ we have that for $(x)\in\prod_{i\in\{1,\ldots,1\}} A_i$ that $[x]_1=\pi_1(x)=x$

Example 5.86. Let $x=(x_1,x_2,x_3)\in\prod_{i\in\{1,\ldots,3\}} A_i$ then

$$\begin{aligned}[x]_3 &= ([x_{|\{1,\ldots 2\}}]_2,\pi_3(x))\\ &= (([x_{|\{1\}}]_1,\pi_2(x)),\pi_3(x))\\ &= ((\pi_1(x),\pi_2(x)),\pi_3(x))\in\bigotimes_{i\in\{1,\ldots,3\}} A_i\end{aligned}$$

Lemma 5.87. *Let* $n\in\mathbb{N}$, $\{A_i\}_{i\in\mathbb{N}}$ *be a family of sets,* $x\in\prod_{i\in\{1,\ldots,n\}} A_i$ *and* $\{B_i\}_{i\in\{1,\ldots,n\}}$ *a family of sets such that* $\forall i\in\{1,\ldots,n\}$ *we have* $B_i\subseteq A_i$ *then* $x\in\prod_{i\in\{1,\ldots,n\}} B_i\Leftrightarrow[x]_n\in\bigotimes_{i\in\{1,\ldots,n\}} B_i$

Proof.

$\Rightarrow$. We prove this by induction, let $\mathcal{S}=\left\{n\in\mathbb{N}|\forall\{A_i\}_{i\in\{1,\ldots,n\}}, x\in\prod_{i\in\{1,\ldots,n\}} A_i, \{B_i\}_{i\in\{1,\ldots,n\}}\right.$ such that $\forall i\in\{1,\ldots,n\}\ B_i$ we have if $x\in\prod_{i\in\{1,\ldots,n\}} B_i$ that $\left.[x]_n\in\bigotimes_{i\in\{1,\ldots,n\}} B_i\right\}$ then we have

$1\in\mathcal{S}$. As $x\in\prod_{i\in\{1,\ldots,1\}} B_i\Rightarrow\pi_1(x)=x(1)\in B_1$ so that $[x]_1=\pi_1(x)\in B_1=\bigotimes_{i\in\{1,\ldots,1\}} B_i$ proving that $1\in\mathcal{S}$.

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$. Assume for $x \in \prod_{i \in \{1,\ldots,n+1\}} A_i$ that $x \in \prod_{i \in \{1,\ldots,n+1\}} B_i$ then using 5.79 we have that $x_{|\{1,\ldots,n\}} \in \prod_{i \in \{1,\ldots,n\}} A_i$ and $x_{|\{1,\ldots,n\}} \in \prod_{i \in \{1,\ldots,n\}} B_i$. Using the induction hypothesis $n \in \mathcal{S}$ we have that $[x_{|\{1,\ldots,n\}}]_n \in \bigotimes_{i \in \{1,\ldots,n\}} B_i$ and as also $\pi_{n+1}(x) = x(n+1) \in B_{n+1}$ it follows that $([x_{|\{1,\ldots,n\}}]_n, \pi_{n+1}(x)) \in \left(\bigotimes_{i \in \{1,\ldots,n\}} B_i\right) \times B_{n+1} = \bigotimes_{i \in \{1,\ldots,n+1\}} B_i$. Hence $x \in \bigotimes_{i \in \{1,\ldots,n+1\}} B_i$ proving that $n+1 \in \mathcal{S}$

$\Leftarrow$. We prove this by induction, let $\mathcal{S} = \left\{n \in \mathbb{N} | \forall \{A_i\}_{i \in \{1,\ldots,n\}}, x \in \prod_{i \in \{1,\ldots,n\}} A_i, \{B_i\}_{i \in \{1,\ldots,n\}} \text{ such that } \forall i \in \{1, \ldots, n\}\ B_i \subseteq A_i \text{ we have if } [x]_n \in \bigotimes_{i \in \{1,\ldots,n\}} B_i \text{ then } x \in \prod_{i \in \{1,\ldots,n\}}\right\}$ then

$1 \in \mathcal{S}$. If $[x]_1 \in \bigotimes_{i \in \{1,\ldots,1\}} B_i$ then $x(1) = \pi_1(x) \in B_1$ so that by 2.83 we have that $x \in \prod_{i \in \{1,\ldots,1\}} B_i$, hence $1 \in \mathcal{S}$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$. Assume for $x \in \prod_{i \in \{1,\ldots,n+1\}} A_i$ that $[x]_{n+1} \in \bigotimes_{i \in \{1,\ldots,n+1\}} B_i = \left(\bigotimes_{i \in \{1,\ldots,n\}} B_i\right) \times B_{n+1}$ then $[x_{|\{1,\ldots,n\}}]_n \in \bigotimes_{i \in \{1,\ldots,n\}} B_i$ and $\pi_{n+1}(x) \in B_{n+1} \Rightarrow x(n+1) \in B_{n+1}$, As $n \in \mathcal{S}$ it follows that $x_{|\{1,\ldots,n\}} \in \prod_{i \in \{1,\ldots,n\}} B_i$. So $\forall i \in \{1, \ldots, n+1\}$ we have $x(i) \in B_i$ from which it follows using 2.83 that $x \in \prod_{i \in \{1,\ldots,n+1\}} B_i$. This proves that $n+1 \in \mathcal{S}$. □

We prove now that $[.]_n$ can be used to define a bijection between $\prod_{i \in \{1,\ldots,n\}} A_i$ and $\bigotimes_{i \in \{1,\ldots,n\}} A_i$.

Theorem 5.88. *Let $n \in \mathbb{N}$, $\{A_i\}_{i \in \{1,\ldots,n\}}$ be a family of sets, then for the function $\mathcal{P}_n: \prod_{i \in \{1,\ldots,n\}} A_i \to \bigotimes_{i \in \{1,\ldots,n\}} A_i$ defined by $\mathcal{P}_n(x) = [x]_n$ we have that*

1. *$\mathcal{P}_n$ is a bijection*
2. *$\forall \{B_i\}_{i \in \{1,\ldots,n\}}$ with $\forall i \in \{1, \ldots, n\}$ $B_i \subseteq A_i$ we have for $\mathcal{P}_n: \prod_{i \in \{1,\ldots,n\}} A_i \to \bigotimes_{i \in \{1,\ldots,n\}} A_i$ that $\mathcal{P}_n\left(\prod_{i \in \{1,\ldots,n\}} B_i\right) = \bigotimes_{i \in \{1,\ldots,n\}} B_i$ (note that this is sensible because of 5.83) [or as $\mathcal{P}_n$ is a bijection $\prod_{i \in \{1,\ldots,n\}} B_i = (\mathcal{P}_n^{-1})\left(\bigotimes_{i \in \{1,\ldots,n\}} B_i\right)$]*

Proof.

1. To prove that $\mathcal{P}_n$ is a bijection we prove that $\mathcal{P}_n$ is a injection and surjection by induction

   injectivity. let $\mathcal{S} = \left\{n \in \mathbb{N} | \text{if } \{A_i\}_{i \in \{1,\ldots,n\}} \text{ is a family of sets then } \mathcal{P}_n: \prod_{i \in \{1,\ldots,n\}} A_i \to \bigotimes_{i \in \{1,\ldots,n\}} A_i \text{ is injectivity}\right\}$ then we have

   $1 \in \mathcal{S}$. then if $\mathcal{P}_1(x) = \mathcal{P}_1(y) \Rightarrow [x]_1 = [y]_1 \Rightarrow \pi_1(x) = \pi_1(y)$ so that $x(1) = y(1)$ proving that the functions $x, y: \{1\} \to A_1$ are the same. Hence $1 \in \mathcal{S}$

$\boldsymbol{n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}}$**.** Assume that $x, y \in \prod_{i\in\{1,\ldots,n+1\}} A_i$ such that $\mathcal{P}_{n+1}(x)=\mathcal{P}_{n+1}(y)$ then

$$\begin{array}{rcl}\mathcal{P}_{n+1}(x)=\mathcal{P}_{n+1}(y) & \Rightarrow & [x]_{n+1}=[x]_{n+1}\\ & \Rightarrow & ([x_{|\{1,\ldots,n\}}]_n, \pi_{n+1}(x)) = ([y_{|\{1,\ldots,n\}}]_n, \pi_{n+1}(y))\\ & \Rightarrow & [x_{|\{1,\ldots,n\}}]_n = [y_{|\{1,\ldots,n\}}]_n \wedge \pi_{n+1}(x)=\pi_{n+1}(y)\\ & \Rightarrow & [x_{|\{1,\ldots,n\}}]_n = [y_{|\{1,\ldots,n\}}]_n \wedge x(n+1)=y(n+1)\\ & \underset{n\in\mathcal{S}}{\Rightarrow} & x_{|\{1,\ldots,n\}} = y_{|\{1,\ldots,n\}} \wedge x(n+1)=y(n+1)\\ & \Rightarrow & \forall i\in\{1,\ldots,n\}\vDash x(i)=y(i)\wedge x(n+1)=y(n+1)\\ & \Rightarrow & x=y\end{array}$$

proving that $\mathcal{P}_{n+1}$ is injective and thus that $n+1\in\mathcal{S}$.

surjectivity. let $\mathcal{S}=\{n\in\mathbb{N}|\text{if } \{A_i\}_{i\in\{1,\ldots,n\}} \text{ is a family of sets then } \mathcal{P}_n\colon \prod_{i\in\{1,\ldots,n\}} A_i \rightarrow \bigotimes_{i\in\{1,\ldots,n\}} A_i \text{ is surjective}\}$ then we have

$\boldsymbol{1\in\mathcal{S}}$**.** then $\bigotimes_{i\in\{1,\ldots,1\}} A_i = A_1$ and if $x\in A_1$ define $\hat{x}\colon\{1\}\rightarrow A_1$ by $\hat{x}(1)=x\in A_1$then $\hat{x}\in\prod_{i\in\{1,\ldots,1\}} A_i$ (see 2.80) and $\mathcal{P}_1(\hat{x})=[\hat{x}]_1=\pi_1(\hat{x})=x$ proving that $\mathcal{P}_1$ is surjective.

$\boldsymbol{n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}}$**.** let $x\in\bigotimes_{i\in\{1,\ldots,n+1\}} A_i$ then as $n+1>1$ we have that $x=(y,z)$ where $y\in\bigotimes_{i\in\{1,\ldots,n\}} A_i$ and $z\in A_{n+1}$. As $n\in\mathcal{S}$ there exists a $\hat{y}\in\prod_{i\in\{1,\ldots,n\}} A_i$ such that $\mathcal{P}_n(\hat{y})=y$. Define now $\hat{x}\in\prod_{i\in\{1,\ldots,n+1\}} A_i$ as $\hat{y}_{+z}$ (see 5.80). Then

$$\begin{array}{rcl}\mathcal{P}_{n+1}(\hat{x}) & = & (\mathcal{P}_n(\hat{x}_{|\{1,\ldots,n\}}), \pi_{n+1}(\hat{x}))\\ & \underset{5.80}{=} & (\mathcal{P}_n(\hat{y}), z)\\ & = & (y,z)\end{array}$$

proving that $\mathcal{P}_{n+1}$ is surjective so that $n+1\in\mathcal{S}$

2. If $x\in\mathcal{P}_n\left(\prod_{i\in\{1,\ldots,n\}} B_i\right)$ then $\exists\hat{x}\in\prod_{i\in\{1,\ldots,n\}} B_i\subseteq\prod_{i\in\{1,\ldots,n\}} A_i$ such $x=[\hat{x}]_n$, using 5.87 we have that $[\bar{x}]_n\in\bigotimes_{i\in\{1,\ldots,n\}} B_i$ proving that $\mathcal{P}\left(\prod_{i\in\{1,\ldots,n\}} B_i\right)\subseteq\bigotimes_{i\in\{1,\ldots,n\}} B_i$. Finally if $x\in\bigotimes_{i\in\{1,\ldots,n\}} B_i$ we have as $\mathcal{P}_n$ is a bijection that there exists a $\hat{x}\in\prod_{i\in\{1,\ldots,n\}} B_i$ such that $x=[\hat{x}]_n$. Applying then 5.87 we have $\hat{x}\in\prod_{i\in\{1,\ldots,n\}} B_i$ so that $x\in\mathcal{P}_n\left(\prod_{i\in\{1,\ldots,n\}} B_i\right)$ proving that $\bigotimes_{i\in\{1,\ldots,n\}} B_i\subseteq\mathcal{P}\left(\prod_{i\in\{1,\ldots,n\}} B_i\right)$. ☐

Lemma 5.89. *Let $n\in\mathbb{N}$, $\{A_i\}_{i\in\{1,\ldots,n\}}$ a finite family of sets then $\bigotimes_{i\in\{1,\ldots,n\}} A_i=\emptyset\Leftrightarrow\exists i\in\{1,\ldots,n\}\vDash A_i=\emptyset$*

Proof. W prove this by induction, so let $\mathcal{S}=\left\{n\in\mathbb{N}|\forall\{A_i\}_{i\in\{1,\ldots,n\}}\text{ we have }\bigotimes_{i\in\{1,\ldots,n\}}A_i=\emptyset\Leftrightarrow\exists i\in\{1,\ldots,n\}\vdash A_i=\emptyset\right\}$ then

$1\in\mathcal{S}$. As $\bigotimes_{i\in\{1,\ldots,1\}}A_i=A_1$ we have $\bigotimes_{i\in\{1,\ldots,1\}}A_i=\emptyset\Leftrightarrow A_1=\emptyset$ proving that $1\in\mathcal{S}$

$n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$. then we have

1. If $\bigotimes_{i\in\{1,\ldots,n+1\}}A_i=\emptyset$ we have

$$\begin{aligned}\bigotimes_{i\in\{1,\ldots,n+1\}}A_i=\emptyset\quad&\underset{n+1>1}{\Leftrightarrow}\quad\left(\bigotimes_{i\in\{1,\ldots,n\}}A_i\right)\times A_{n+1}=\emptyset\\&\underset{1.43}{\Leftrightarrow}\quad\left(\bigotimes_{i\in\{1,\ldots,n\}}A_i\right)=\emptyset\vee A_{n+1}=\emptyset\\&\underset{n\in\mathcal{S}}{\Leftrightarrow}\quad(\exists i\in\{1,\ldots,n\}\vDash A_i=\emptyset)\vee A_{n+1}=\emptyset\\&\Leftrightarrow\quad\exists i\in\{1,\ldots,n+1\}\vDash A_i=\emptyset\end{aligned}$$

□

Using the above bijection we can define the projection function.

Definition 5.90. *Let $n\in\mathbb{N},\{A_i\}_{i\in\{1,\ldots,n\}}$ a finite family of sets, $i\in\{1,\ldots,n\}$ then we define $\pi_i^{\times}:\bigotimes_{j\in\{1,\ldots,n\}}A_j\rightarrow A_i$ by $\pi_i^{\times}=\pi_i\circ\mathcal{P}_n^{-1}$*

Example 5.91. For the special case $\{A_i\}_{i\in\{1,\ldots,1\}}$ we have that $\pi_1^{\times}=\mathbb{1}_{A_1}$

Proof. Let $x\in A_1\underset{\text{def}}{=}\bigotimes_{i\in\{1,\ldots,1\}}A_i$ and define $x'\in\prod_{i\in\{1,\ldots,1\}}A_i$ by $x'=\{(1,x)\}$ then $\mathcal{P}_1(x')=[x']=\pi_1(x')=x$ so that $\mathcal{P}_1^{-1}(x)=x'$ hence $\pi_1^{\times}(x)=\pi_1(\mathcal{P}_1^{-1}(x))=\pi_1(x')=x$ □

Theorem 5.92. *Let $n\in\mathbb{N}$, $\{A_i\}_{i\in\{1,\ldots,n\}}$ a finite family of sets then we have*

1. *$\forall x,y\in\bigotimes_{i\in\{1,\ldots,n\}}A_i$ we have $x=y\Leftrightarrow\forall i\in\{1,\ldots,n\}$ we have $\pi_i^{+}(x)=\pi_i^{+}(y)$*
2. *If $\forall i\in\{1,\ldots,n\}\;A_i\neq\emptyset$ then $\forall i\in\{1,\ldots,n\}$ we have that $\pi_i^{\times}:\bigotimes_{i\in\{1,\ldots,n\}}A_i\rightarrow A_i$ is a surjection*
3. *Let $\{B_i\}_{i\in\{1,\ldots,n\}}$ another finite family of sets then if $\forall x\in\bigotimes_{i\in\{1,\ldots,n\}}A_i$ we have that $\forall i\in\{1,\ldots,n\}\;\pi_i^{\times}(x)\in B_i$ then $\bigotimes_{i\in\{1,\ldots,n\}}A_i\subseteq\bigotimes_{i\in\{1,\ldots,n\}}B_i$*

Proof.

1. If $x,y\in\bigotimes_{i\in\{1,\ldots,n\}}A_i$ then

$$\begin{aligned}x=y\quad&\underset{\mathcal{P}_n\text{ is a bijection}}{\Leftrightarrow}\quad\mathcal{P}_n^{-1}(x)=\mathcal{P}_n^{-1}(y)\\&\Leftrightarrow\quad\forall i\in\{1,\ldots,n\}\text{ we have }(\mathcal{P}_n^{-1}(x))(i)=(\mathcal{P}_n^{-1}(y))(i)\\&\Leftrightarrow\quad\forall i\in\{1,\ldots,n\}\text{ we have }\pi_i(\mathcal{P}_n^{-1}(x))=\pi_i(\mathcal{P}_n^{-1}(y))\\&\Leftrightarrow\quad\forall i\in\{1,\ldots,n\}\text{ we have }\pi_i^{\times}(x)=\pi_i^{\times}(y)\end{aligned}$$

2. As $\forall i \in \{1, \ldots, n\}$ $A_i \neq \emptyset$ we have by 2.89 that $\forall i \in \{1, \ldots, n\}$ π_i is a surjection. Take $i \in \{1, \ldots, n\}$ and $x \in A_i$ then there exists a $\hat{x} \in \prod_{i \in \{1, \ldots, n\}} A_i$ such that

$$\begin{array}{rcl} x & = & \pi_1(\hat{x}) \\ & \underset{\text{bijective}}{=} & \pi_i((\mathcal{P}_n^{-1})(\mathcal{P}_n(\hat{x}))) \\ & = & \pi_i^{\times}(\mathcal{P}_n(\hat{x})) \end{array}$$

which as $\mathcal{P}_n(x) \in \bigotimes_{i \in \{1, \ldots, n\}} A_i$ proves that $\pi_i^{\times}$ is a surjection.

3. We have to consider two cases

$\exists i \in \{1, \ldots, n\} \vdash A_i = \emptyset$. then using 5.89 we have that $\bigotimes_{i \in \{1, \ldots, n\}} A_i = \emptyset$ so that

$$\bigotimes_{i \in \{1, \ldots, n\}} A_i \subseteq \bigotimes_{i \in \{1, \ldots, n\}} B_i.$$

$\forall i \in \{1, \ldots, n\} \vDash A_i \neq \emptyset$. Take $i \in \{1, \ldots, n\}$ then by (2) $\pi_i^{\times}$ is a surjection, so if $x \in A_i$ there exists a $\hat{x} \in \bigotimes_{i \in \{1, \ldots, n\}} A_i$ such that $x = \pi_i^{\times}(\hat{x})$. As by the hypothesis we have that $\pi_i^{\times}(\hat{x}) \in B_i$ it follows that $x \in B_i$. Hence $\forall i \in \{1, \ldots, n\}$ we have $A_i \subseteq B_i$ and using 5.83 it follows that

$$\bigotimes_{i \in \{1, \ldots, n\}} A_i \subseteq \bigotimes_{i \in \{1, \ldots, n\}} B_i. \qquad \square$$

In the next theorem we examine the product of product of sets.

Theorem 5.93. *Let $n, k \in \mathbb{N}$ such that $1 \leqslant k < n$ and $\{A_i\}_{i \in \{1, \ldots, n\}}$ a finite family of sets then there exists a bijection $\beta: \bigotimes_{i \in \{1, \ldots, n\}} A_i \rightarrow \left(\bigotimes_{i \in \{1, \ldots, k\}} A_i\right) \times \left(\bigotimes_{i \in \{1, \ldots, n-k\}} A_{k+i}\right)$ such that given $x \in \bigotimes_{i \in \{1, \ldots, n\}} A_i$ we have*

$$\forall i \in \{1, \ldots, k\}\ \pi_{n,i}^{\times}(x) = \pi_{k.i}(\pi_1(\beta(x)))$$

and

$$\forall i \in \{k+1, \ldots, n\}\ \pi_{n,i}^{\times}(x) = \pi_{n-k,i-k}(\pi_2(\beta(x)))$$

where $\pi_{n,i}^{\times}$ is the projection function between $\bigotimes_{i \in \{1, \ldots, n\}} A_i$ and A_i, $\pi_{k,i}^{\times}$ is the projection function between $\bigotimes_{i \in \{1, \ldots, k\}} A_i$ and A_i, $\pi_{n-k,i}^{\times}$ is the projection function between $\bigotimes_{i \in \{1, \ldots, n-k\}} A_{i+k}$ and A_{i+k}, π_1 is the projection function between $\left(\bigotimes_{i \in \{1, \ldots, k\}} A_i\right) \times \left(\bigotimes_{i \in \{1, \ldots, n-k\}} A_{k+i}\right)$ and $\bigotimes_{i \in \{1, \ldots, k\}} A_i$ and π_2 is the projection function between $\left(\bigotimes_{i \in \{1, \ldots, k\}} A_i\right) \times \left(\bigotimes_{i \in \{1, \ldots, n-k\}} A_{k+i}\right)$ and $\bigotimes_{i \in \{1, \ldots, n-k\}} A_{k+i}$.

Proof. We use the following bijections and functions in the constructing of the bijection β:

$\alpha: \prod_{i \in \{1, \ldots, n\}} A_i \rightarrow \left(\prod_{i \in \{1, \ldots, k\}} A_i\right) \times \left(\prod_{i \in \{k+1, \ldots, n\}} A_i\right)$ where $\alpha(x) = (x_{|\{1, \ldots, k\}}, x_{|\{k+1, \ldots, n\}})$

which is a bijection because 5.79. As $\rho: \{k+1, \ldots, n\} \to \{1, \ldots, n-k\}$ defined by $\rho(i) = i - k$ is a bijection with inverse $\rho^{-1}: \{1, \ldots, n-k\} \to \{k+1, \ldots, n\}$ defined by $(\rho^{-1})(i) = i + k$. Further we have by 2.81 that

$$\lambda: \prod_{i \in \{k+1, \ldots, n\}} A_i \to \prod_{i \in \{1, \ldots, n-k\}} A_{i+k} \text{ defined by } \lambda(x) = x \circ \rho^{-1}$$

is a bijection. Finally we need the following

$$\mathcal{P}_n: \prod_{i \in \{1, \ldots, n\}} A_i \to \bigotimes_{i \in \{1, \ldots, n\}} A_i \text{ (see 5.88)}$$

$$\mathcal{P}_k: \prod_{i \in \{1, \ldots, k\}} A_i \to \bigotimes_{i \in \{1, \ldots, k\}} A_i \text{ (see 5.88)}$$

$$\mathcal{P}_{n-k}: \prod_{i \in \{1, \ldots, n-k\}} A_i \to \bigotimes_{i \in \{1, \ldots, n-k\}} A_{i+k}$$

We are ready now to define

$$\beta: \bigotimes_{i \in \{1, \ldots, n\}} A_i \to \left(\bigotimes_{i \in \{1, \ldots, k\}} A_i \right) \times \left(\bigotimes_{i \in \{1, \ldots, n-k\}} A_{k+i} \right)$$

by

$$\beta(x) = (\mathcal{P}_k(\pi_1(\alpha(\mathcal{P}_n^{-1}(x)))), \mathcal{P}_{n-k}(\lambda(\pi_2(\alpha(\mathcal{P}_n^{-1}(x))))))$$

Then we have for $x \in \bigotimes_{i \in \{1, \ldots, n\}} A_i$ that for $i \in \{1, \ldots, k\}$

$$\begin{aligned}
\pi^{\times}_{k,i}(\pi_1(\beta(x))) &= \pi^{\times}_{k,i}(\mathcal{P}_k(\pi_1(\alpha(\mathcal{P}_n^{-1}(x))))) \\
&= \pi_{k,i}(\mathcal{P}_k^{-1}(\mathcal{P}_k(\pi_1(\alpha(\mathcal{P}_n^{-1}(x)))))) \\
&= \pi_{k,i}(\pi_1(\alpha(\mathcal{P}_n^{-1}(x)))) \\
&= \pi_{k,i}(\mathcal{P}_n^{-1}(x)_{|\{1, \ldots, k\}}) \\
&= \mathcal{P}_n^{-1}(x)_{|\{1, \ldots, k\}}(i) \\
&= \mathcal{P}_n^{-1}(x)(i) && (5.5) \\
&= \pi_{n,i}(\mathcal{P}_n^{-1}(x)) \\
&= \pi^{\times}_{n,i}(x) && (5.6)
\end{aligned}$$

and for $i \in \{k+1, \ldots, n-k\}$

$$\begin{aligned}
\pi^{\times}_{n-k,i-k}(\pi_2(\beta(x))) &= \pi^{\times}_{n-k,i-k}(\mathcal{P}_{n-k}(\lambda(\pi_2(\alpha(\mathcal{P}_n^{-1}(x)))))) \\
&= \pi_{n-k,i-k}(\mathcal{P}_{n-k}^{-1}(\mathcal{P}_{n-k}(\lambda(\pi_2(\alpha(\mathcal{P}_n^{-1}(x))))))) \\
&= \pi_{n-k,i-k}(\lambda(\pi_2(\alpha(\mathcal{P}_n^{-1}(x))))) \\
&= \pi_{n-k,i-k}(\lambda(\mathcal{P}_n^{-1}(x)_{|\{k+1, \ldots, n\}})) \\
&= \lambda(\mathcal{P}_n^{-1}(x)_{|\{k+1, \ldots, n\}})(i-k) \\
&= \mathcal{P}_n^{-1}(x)_{|\{k+1, \ldots, n\}}(i) \\
&= \mathcal{P}_n^{-1}(x)(i) && (5.7) \\
&= \pi_{n,i}(\mathcal{P}_n^{-1}(x)) \\
&= \pi^{\times}_{n,i}(x) && (5.8)
\end{aligned}$$

The above proves part of the theorem, for the rest we have to prove that β is a bijection:

injectivity. Let $x, y \in \bigotimes_{i\in\{1,\ldots,\}} A_i$ such that $\beta(x) = \beta(y)$ then we have $\forall i \in \{1, \ldots, n\}$

$$\begin{array}{rcl} \pi^{\times}_{k,i}(\pi_1(\beta(x))) = \pi^{\times}_{k,i}(\pi_1(\beta(y))) & \underset{5.5,5.7}{\Rightarrow} & \mathcal{P}_n^{-1}(x)(i) = \mathcal{P}_n^{-1}(x)(i) \\ & \Rightarrow & \mathcal{P}_n^{-1}(x) = \mathcal{P}_n^{-1}(x) \\ & \underset{\mathcal{P}_n \text{ is bijective}}{\Rightarrow} & x = y \end{array}$$

proving injectivity.

surjectivity. Let $(x, y) \in \left(\bigotimes_{i\in\{1,\ldots,k\}} A_i\right) \times \left(\bigotimes_{i\in\{k+1,\ldots,n\}} A_i\right)$ take then

$$z = \mathcal{P}_n(\alpha^{-1}((\mathcal{P}_k^{-1}(x)), \lambda^{-1}(\mathcal{P}_{n-k}^{-1}(y))))$$

then we have

$$\begin{array}{rcl} \mathcal{P}_k(\pi_1(\alpha(\mathcal{P}_n^{-1}(z)))) & = & \mathcal{P}_k(\pi_1(\alpha(\mathcal{P}_n^{-1}(\mathcal{P}_n(\alpha^{-1}((\mathcal{P}_k^{-1}(x)), \lambda^{-1}(\mathcal{P}_{n-k}^{-1}(y)))))))) \\ & = & \mathcal{P}_k(\pi_1(\alpha(\alpha^{-1}((\mathcal{P}_k^{-1}(x)), \lambda^{-1}(\mathcal{P}_{n-k}^{-1}(y)))))) \\ & = & \mathcal{P}_k(\pi_1((\mathcal{P}_k^{-1}(x)), \lambda^{-1}(\mathcal{P}_{n-k}^{-1}(y)))) \\ & = & \mathcal{P}_k(\mathcal{P}_k^{-1}(x)) \\ & = & x \end{array}$$

and

$$\begin{array}{rcl} \mathcal{P}_{n-k}(\lambda(\pi_2(\alpha(\mathcal{P}_n^{-1}(z))))) & = & \mathcal{P}_{n-k}(\lambda(\pi_2(\alpha(\mathcal{P}_n^{-1}(\mathcal{P}_n(\alpha^{-1}((\mathcal{P}_k^{-1}(x)), \\ & & \lambda^{-1}(\mathcal{P}_{n-k}^{-1}(y))))))))) \\ & = & \mathcal{P}_{n-k}(\lambda(\pi_2(\alpha(\alpha^{-1}((\mathcal{P}_k^{-1}(x)), \\ & & \lambda^{-1}(\mathcal{P}_{n-k}^{-1}(y))))))) \\ & = & \mathcal{P}_{n-k}(\lambda(\pi_2((\mathcal{P}_k^{-1}(x)), \lambda^{-1}(\mathcal{P}_{n-k}^{-1}(y))))) \\ & = & \mathcal{P}_{n-k}(\lambda(\lambda^{-1}(\mathcal{P}_{n-k}^{-1}(y)))) \\ & = & \mathcal{P}_{n-k}(\mathcal{P}_{n-k}^{-1}(y)) \\ & = & y \end{array}$$

so that using the above

$$\begin{array}{rcl} \beta(z) & = & (\mathcal{P}_k(\pi_1(\alpha(\mathcal{P}_n^{-1}(z)))), \mathcal{P}_{n-k}(\lambda(\pi_2(\alpha(\mathcal{P}_n^{-1}(z)))))) \\ & = & (x, y) \end{array}$$

proving that β is surjective. □

Next we prove that $\bigotimes_{i\in\{1,\ldots,n\}} A_i$ has the same properties as $\prod_{i\in\{1,\ldots,n\}} A_i$ (more specific 2.86,2.91,2.93,2.95,2.96,2.97)

Theorem 5.94. *Let $n \in \mathbb{N}$, $\{A_i\}_{i\in\{1,\ldots,n\}}, \{B_i\}_{i\in\{1,\ldots,n\}}$ a finite family of sets with $\forall i \in \{1, \ldots, n\}$ $A_i \neq \emptyset$ (which by the previous theorem means that $\bigotimes_{i\in\{1,\ldots,n\}} A_i \neq \emptyset$) then $\bigotimes_{i\in\{1,\ldots,n\}} A_i \subseteq \bigotimes_{i\in\{1,\ldots,n\}} B_i \Leftrightarrow A_i \subseteq B_i$*

Proof. We prove this by induction, so let $\mathcal{S}=\left\{n\in\mathbb{N}|\{A_i\}_{i\in\{1,\ldots,n\}},\{B_i\}_{i\in\{1,\ldots,n\}}\right.$ with $\exists i\in\{1,\ldots,n\}\vdash A_i\neq\emptyset$ then $\left.\bigotimes_{i\in\{1,\ldots,n\}}A_i\subseteq\bigotimes_{i\in\{1,\ldots,n\}}B_i\Leftrightarrow A_i\subseteq B_i\right\}$ then we have

$1\in\mathcal{S}$. As $\bigotimes_{i\in\{1,\ldots,1\}}A_i=A_1$ and $\bigotimes_{i\in\{1,\ldots,1\}}B_i=B_1$ we have $\bigotimes_{i\in\{1,\ldots,1\}}A_i\subseteq\bigotimes_{i\in\{1,\ldots,1\}}B_i$ if and only $A_1\subseteq B_1$ proving that $1\in\mathcal{S}/$

$n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$. Take $\bigotimes_{i\in\{1,\ldots,n+1\}}A_i$, $\bigotimes_{i\in\{1,\ldots,n+1\}}B_i$ then

$\Rightarrow$. If $\forall i\in\{1,\ldots,n+1\}$ we have that $A_i\neq\emptyset$ we have using the theorem 5.89 that

$$\bigotimes_{i\in\{1,\ldots,n\}}A_i\neq\emptyset\wedge A_{n+1}\neq\emptyset \tag{5.9}$$

$$\begin{aligned}
\bigotimes_{i\in\{1,\ldots,n+1\}}A_i\subseteq\bigotimes_{i\in\{1,\ldots,n+1\}}B_i &\underset{1<n+1}{\Rightarrow} \left(\bigotimes_{i\in\{1,\ldots,n\}}A_i\right)\times A_{n+1}\subseteq\left(\bigotimes_{i\in\{1,\ldots,n\}}B_i\right)\times B_{n+1}\\
&\underset{5.9\text{ and }1.45}{\Rightarrow} \bigotimes_{i\in\{1,\ldots,n\}}A_i\subseteq\bigotimes_{i\in\{1,\ldots,n\}}B_i\wedge A_{n+1}\subseteq B_{n+1}\\
&\underset{n\in\mathcal{S}}{\Rightarrow} (\forall i\in\{1,\ldots,n\}\vDash A_i\subseteq B_i)\wedge A_{n+1}\subseteq B_{n+1}\\
&\Rightarrow \forall i\in\{1,\ldots,n+1\}\vDash A_i\subseteq B_i
\end{aligned}$$

$\Leftarrow$. Using 5.83 we have if $\forall i\in\{1,\ldots,n+1\}\vDash A_i\subseteq B_i$ then $\bigotimes_{i\in\{1,\ldots,n+1\}}A_i\subseteq\bigotimes_{i\in\{1,\ldots,n+1\}}B_i$.

So we have proved that $\bigotimes_{i\in\{1,\ldots,n+1\}}A_i\subseteq\bigotimes_{i\in\{1,\ldots,n+1\}}B_i\Leftrightarrow\forall i\in\{1,\ldots,n+1\}\vDash A_i\subseteq B_i$ from which we conclude that $n+1\in\mathcal{S}$ □

Corollary 5.95. *Let $n\in\mathbb{N}$, $\{A_i\}_{i\in\{1,\ldots,n\}},\{B_i\}_{i\in\{1,\ldots,n\}}$ finite families of sets with $\forall i\in\{1,\ldots,n\}$ $A_i\neq\emptyset$ then $\bigotimes_{i\in\{1,\ldots,n\}}A_i=\bigotimes_{i\in\{1,\ldots,n\}}B_i\Leftrightarrow\forall i\in\{1,\ldots,n\}$ we have $A_i=B_1$*

Proof.

$\Rightarrow$. As $\forall i\in\{1,\ldots,n\}$ $A_i\neq\emptyset$ we have by 5.89 that $\emptyset\neq\bigotimes_{i\in\{1,\ldots,n\}}A_i=\bigotimes_{i\in\{1,\ldots,n\}}B_i$ or applying 5.89 again we have $\forall i\in\{1,\ldots,n\}$ $B_i\neq\emptyset$. Then we have

$$\begin{aligned}
\bigotimes_{i\in\{1,\ldots,n\}}A_i=\bigotimes_{i\in\{1,\ldots,n\}}B_i &\Rightarrow \bigotimes_{i\in\{1,\ldots,n\}}A_i\subseteq\bigotimes_{i\in\{1,\ldots,n\}}B_i\\
&\underset{5.94}{\Rightarrow} \forall i\in\{1,\ldots,n\}\vDash A_i\subseteq B_i
\end{aligned}$$

and similar

$$\bigotimes_{i\in\{1,\ldots,n\}} A_i = \bigotimes_{i\in\{1,\ldots,n\}} B_i \Rightarrow \bigotimes_{i\in\{1,\ldots,n\}} B_i \subseteq \bigotimes_{i\in\{1,\ldots,n\}} A_i$$
$$\underset{5.94}{\Rightarrow} \forall i\in\{1,\ldots,n\} \vDash B_i \subseteq A_i$$

so we have $\forall i\in\{1,\ldots,n\}$ that $A_i = B_i$

$\Leftarrow$. This is trivial □

Theorem 5.96. *Let $n\in\mathbb{N}$, $\{A_i\}_{i\in\{1,\ldots,n\}}, \{B_i\}_{i\in\{1,\ldots,n\}}$ finite families of sets then $\left(\bigotimes_{i\in\{1,\ldots,n\}} A_i\right)\bigcap\left(\bigotimes_{i\in\{1,\ldots,n\}} B_i\right) = \bigotimes_{i\in\{1,\ldots,n\}}\left(A_i \bigcap B_i\right)$*

Proof. First as $\forall i\in\{1,\ldots,n\}$ we have that $A_i\bigcap B_i \subseteq A_i, B_i$ we have using 5.83 that $\bigotimes_{i\in\{1,\ldots,n\}}(A_i\bigcap B_i)\subseteq \bigotimes_{i\in\{1,\ldots,n\}} A_i$, $\bigotimes_{i\in\{1,\ldots,n\}} B_i$ proving that

$$\bigotimes_{i\in\{1,\ldots,n\}}\left(A_i\bigcap B_i\right)\subseteq\left(\bigotimes_{i\in\{1,\ldots,n\}} A_i\right)\bigcap\left(\bigotimes_{i\in\{1,\ldots,n\}} B_i\right) \tag{5.10}$$

For the opposite inclusion let $x\in\left(\bigotimes_{i\in\{1,\ldots,n\}} A_i\right)\bigcap\left(\bigotimes_{i\in\{1,\ldots,n\}} B_i\right)$ then $x\in \bigotimes_{i\in\{1,\ldots,n\}} A_i \wedge x\in\bigotimes_{i\in\{1,\ldots,n\}} B_i$. Using 5.92 (3) we have that $\forall i\in\{1,\ldots,n\}$ $\pi_i^{\times}(x)\in A_i$ and $\pi_i^{\times}(x)\in B_i$ so that $\pi_i^{\times}(x)\in A_i\bigcap B_i$. Finally using 5.92 (3) again we have $\left(\bigotimes_{i\in\{1,\ldots,n\}} A_i\right)\bigcap\left(\bigotimes_{i\in\{1,\ldots,n\}} B_i\right)\subseteq\bigotimes_{i\in\{1,\ldots,n\}}(A_i\bigcap B_i)$ which together with 5.10 gives

$$\left(\bigotimes_{i\in\{1,\ldots,n\}} A_i\right)\bigcap\left(\bigotimes_{i\in\{1,\ldots,n\}} B_i\right)=\bigotimes_{i\in\{1,\ldots,n\}}\left(A_i\bigcap B_i\right)$$

□

Actually we can generalize the above to a arbitrary collection of family of sets with almost the same proof.

Theorem 5.97. *Let $n\in\mathbb{N}$, I be a set, and $\{\{A_{i,j}\}_{j\in\{1,\ldots,n\}}\}_{i\in I}$ a family of finite families of sets then $\bigcap_{i\in I}\left(\bigotimes_{j\in\{1,\ldots,n\}} A_{i,j}\right)=\bigotimes_{j\in\{1,\ldots,n\}}\left(\bigcap_{i\in I} A_{i,j}\right)$*

Proof. First $\forall i\in I$ we have $\forall j\in\{1,\ldots,n\}$ that $\bigcap_{k\in I} A_{k,j}\subseteq A_{i,j}$ so using 5.83 it follows that $\bigotimes_{j\in\{1,\ldots,n\}}\left(\bigcap_{k\in I} A_{k,j}\right)\subseteq\bigotimes_{k\in\{1,\ldots,n\}} A_{i,k}$ proving that

$$\bigotimes_{j\in\{1,\ldots,n\}}\left(\bigcap_{i\in I} A_{i,j}\right)\subseteq\bigcap_{i\in I}\left(\bigotimes_{j\in\{1,\ldots,n\}} A_{i,j}\right) \tag{5.11}$$

For the opposite inclusion let $x\in\bigcap_{i\in I}\left(\bigotimes_{j\in\{1,\ldots,n\}} A_{i,j}\right)$ then $\forall i\in I$ $x\in \bigotimes_{j\in\{1,\ldots,n\}} A_{i,j}$, using 5.92 (3) we have that $\forall j\in\{1,\ldots,n\}$ $\pi_j^{\times}(x)\in A_{i,j}$, so that $\pi_j^{\times}(x)\in\bigcap_{i\in I} A_{i,j}$. Finally using 5.92 (3) again we have $\bigcap_{i\in I}\left(\bigotimes_{j\in\{1,\ldots,n\}} A_{i,j}\right)\subseteq \bigotimes_{j\in\{1,\ldots,n\}}\left(\bigcap_{i\in I} A_{i,j}\right)$ which together with 5.11 gives

$$\bigcap_{i\in I}\left(\bigotimes_{j\in\{1,\ldots,n\}} A_{i,j}\right)=\bigotimes_{j\in\{1,\ldots,n\}}\left(\bigcap_{i\in I} A_{i,j}\right)$$

□

Theorem 5.98. *Let $n\in\mathbb{N}$, $\{A_i\}_{i\in\{1,\ldots,n\}}$, $\{B_i\}_{i\in\{1,\ldots,n\}}$ be finite family of sets then* $\left(\bigotimes_{i\in\{1,\ldots,n\}} A_i\right)\bigcup\left(\bigotimes_{i\in\{1,\ldots,n\}} B_i\right)\subseteq\bigotimes_{i\in\{1,\ldots,n\}}(A_i\bigcup B_i)$

Proof. As $\forall i\in\mathbb{N}$ we have $A_i\subseteq A_i\bigcup B_i$ and $B_i\subseteq A_i\bigcup B_i$ we have by 5.83 that $\bigotimes_{i\in\{1,\ldots,n\}} A_i\subseteq\bigotimes_{i\in\{1,\ldots,n\}}(A_i\bigcup B_i)$ and $\bigotimes_{i\in\{1,\ldots,n\}} B_i\subseteq\bigotimes_{i\in\{1,\ldots,n\}}(A_i\bigcup B_i)$ so that $\left(\bigotimes_{i\in\{1,\ldots,n\}} A_i\right)\bigcup\left(\bigotimes_{i\in\{1,\ldots,n\}} B_i\right)\subseteq\bigotimes_{i\in\{1,\ldots,n\}}(A_i\bigcup B_i)$ □

Note 5.99. The opposite inclusion is not true for example if $A_1=\{1\}, A_2=\{2\}$, $B_1=\{3\}$and $B_2=\{4\}$ then $(A_1\times A_2)\bigcup(B_1\times B_2)=\{(1,2),\{3,4)\}$ but $(A_1\bigcup B_1)\times(A_2\bigcup B_2)=\{(1,3),(1,4),(2,3),(2,4)\}$

5.6.1 Finite product of sets and denumerability and countability

Theorem 5.100. *Let $\{A_i\}_{i\in\{1,\ldots,n\}}$ be a finite family of sets then we have*

1. *If $\forall i\in\{1,\ldots n\}\vDash A_i$ is finite then $\bigotimes_{i\in\{1,\ldots,n\}} A_i$ is finite and $\#\left(\bigotimes_{i\in\{1,\ldots,n\}} A_i\right)=\prod_{i=1}^n\#(A_i)$*
2. *If $\forall i\in\{1,\ldots,n\}\vDash A_i$ is denumerable then $\bigotimes_{i\in\{1,\ldots,n\}} A_i$ is denumerable*

Proof. We prove this by induction

1. Define $\mathcal{S}=\left\{n\in\mathbb{N}|\text{ If }\{A_i\}_{\{1,\ldots,n\}}\text{ is a finite family of finite sets then }\bigotimes_{i\in\{1,\ldots,n\}} A_i\text{ is finite and }\#\left(\bigotimes_{i\in\{1,\ldots,n\}} A_i\right)=\prod_{i=1}^n\#(A_i)\right\}$ then we have

   $1\in\mathcal{S}$. If $\{A_i\}_{i\in\{1,\ldots,1\}}$ is finite family of finite sets then $\bigotimes_{i\in\{1,\ldots,1\}} A_i=A_1$ is finite and $\prod_{i=1}^1\#(A_i)=\#(A_1)=\#\left(\bigotimes_{i\in\{1,\ldots,n\}} A_i\right)$ proving that $1\in S$

   $n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$. If $\{A_i\}_{i\in\{1,\ldots,n+1\}}$ is a finite family of finite sets then as $n\in\mathcal{S}$ we have that $\bigotimes_{i\in\{1,\ldots,n\}} A_i$ is finite and $\#\left(\bigotimes_{i\in\{1,\ldots,n\}} A_i\right)=\prod_{i=1}^n\#(A_i)$. Now $\bigotimes_{i\in\{1,\ldots,n+1\}} A_i=\left(\bigotimes_{i\in\{1,\ldots,n\}} A_i\right)\bigotimes A_{n+1}$ which is by 5.44 finite and $\#\left(\bigotimes_{i\in\{1,\ldots,n+1\}} A_i\right)=\#\left(\left(\bigotimes_{i\in\{1,\ldots,n\}} A_i\right)\bigotimes A_{n+1}\right)=\#\left(\bigotimes_{i\in\{1,\ldots,n\}} A_i\right)\cdot\#(A_{n+1})=\left(\prod_{i=1}^n\#(A_i)\right)\cdot\#(A_{n+1})=\prod_{i=1}^{n+1}\#(A_i)$ showing that $n+1\in\mathcal{S}$

2. Define $\mathcal{S}=\left\{n\in\mathbb{N}|\text{ If }\{A_i\}_{\{1,\ldots,n\}}\text{ is a finite family of denumerable sets then }\bigotimes_{i\in\{1,\ldots,n\}} A_i\text{ is denumerable}\right\}$ then we have

   $1\in\mathcal{S}$. If $\{A_i\}_{i\in\{1,\ldots,1\}}$ is finite family of denumerable sets then $\bigotimes_{i\in\{1,\ldots,1\}} A_i=A_1$ is denumerable proving that $1\in S$

$\boldsymbol{n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}}$. If $\{A_i\}_{i\in\{1,\ldots,n+1\}}$ is a finite family of denumerable sets then as $n \in \mathcal{S}$ we have that $\bigotimes_{i\in\{1,\ldots,n\}} A_i$ is denumerable. Now $\bigotimes_{i\in\{1,\ldots,n+1\}} A_i = \left(\bigotimes_{i\in\{1,\ldots,n\}} A_i\right) \bigotimes A_{n+1}$ which is by 5.59 denumerable showing that $n+1 \in \mathcal{S}$ □

Theorem 5.101. *Let $\{A_i\}_{i\in\{1,\ldots.,n\}}$ be a finite family of sets then we have*

1. *If $\forall i \in \{1, \ldots, n\} \vDash A_i$ is finite then $\prod_{i\in\{1,\ldots,n\}} A_i$ is finite and $\#\left(\prod_{i\in\{1,\ldots,n\}} A_i\right) = \prod_{i=1}^{n} \#(A_i)$*
2. *If $\forall i \in \{1, \ldots, n\} \vDash A_i$ is denumerable then $\prod_{i\in\{1,\ldots,n\}} A_i$ is denumerable*

Proof.

1. Let $\{A_i\}_{i\in\{1,\ldots,n\}}$ be a finite family of finite sets then using the previous lemma we have that $\bigotimes_{i\in\{1,\ldots,n\}} A_i$ is finite with $\#\left(\bigotimes_{i\in\{1,\ldots,n\}} A_i\right) = \prod_{i=1}^{n} \#(A_i)$. As by 5.88 $\bigotimes_{i\in\{1,\ldots,n\}} A_i$ is bijective to $\prod_{i\in\{1,\ldots,n\}} A_i$ we have that $\prod_{i\in\{1,\ldots,n\}} A_i$ is finite and $\#\left(\prod_{i\in\{1,\ldots,n\}} A_i\right) = \prod_{i=1}^{n} \#(A_i)$.
2. Let $\{A_i\}_{i\in\{1,\ldots,n\}}$ be a finite family of denumerable sets then using the previous lemma we have that $\bigotimes_{i\in\{1,\ldots,n\}} A_i$ is denumerable. As by 5.88 $\bigotimes_{i\in\{1,\ldots,n\}} A_i$ is bijective to $\prod_{i\in\{1,\ldots,n\}} A_i$ we have that $\prod_{i\in\{1,\ldots,n\}} A_i$ is denumerable. □

5.6.2 Notation conventions

In the above we have proved that $\prod_{i\in I} A_i$ and $\bigotimes_{i\in\{1,\ldots,n\}} A_i$ have similar properties and also that if $I = \{1, \ldots, n\}$ we can set up a bijection between $\prod_{i\in\{1,\ldots,n\}} A_i$ and $\bigotimes_{i\in\{1,\ldots,n\}} A_i$. $\prod_{i\in I} A_i$ is of course the most general definition because it also work for non finite families of sets. However there is not a easy way to use a induction argument if $I = \{1, \ldots, n\}$ as $\prod_{i\in\{1,\ldots,n\}} A_i$ is not defined using recursion, this is the big benefit of using $\bigotimes_{i\in\{1,\ldots,n\}} A_i$. Of course we can use the bijection $\mathcal{P}_n: \prod_{i\in\{1,\ldots,n\}} A_i \to \bigotimes_{i\in\{1,\ldots,n\}} A_i$ to prove a theorem in $\bigotimes_{i\in\{1,\ldots,n\}} A_i$ and then the inverse bijection to state the theorem for $\prod_{i\in\{1,\ldots,n\}} A_i$ but this a little cumbersome, so in most cases we work in the finite case only with $\bigotimes_{i\in\{1,\ldots,n\}A_i}$. To avoid double notation and also to be consistent with existing literature we introduce the following notation conventions.

Notation 5.102. (Conventions for product of sets)

1. Unless mentioned otherwise we use the symbol $\prod_{i\in\{1,\ldots,n\}} A_i$ to mean $\bigotimes_{i\in\{1,\ldots,n\}} A_i$ and if I is not finite $\prod_{i\in I} A_i$ is equal to the generalized product.
2. For $\prod_{i\in\{1,\ldots,n\}} A_i$ [which is using (1) $\bigotimes_{i\in\{1,\ldots,n\}} A_i$] we use π_i to mean $\pi_i^{\times}$
3. For $\prod_{i\in I} A_i$ [the general not finite case] π_i is the general projection function (see 2.87

4. If $x \in \prod_{i \in \{1,\ldots,n\}} A_i$ [which is using (1) $\bigotimes_{i \in \{1,\ldots,n\}} A_i$] then as by 5.92 (1) x is uniquely defined by the values of $\pi_i^{\times}(x) \underset{(2)}{=} \pi_i(x)$ we note x_i instead of $\pi_i(x) \underset{(3)}{=} \pi_i^{\times}(x)$. Further we use the notation $x = (x_1, \ldots, x_n)$ to mean that $\forall i \in \{1, \ldots, n\}$ $\pi_i(x) = x_i$.
5. Using the above notation we have that $\prod_{i \in \{1\ldots,1\}} A_i = A_1$ and if $\forall x \in \prod_{i \in \{1\ldots,1\}} A_i = A_1$ we have $x_1 = \pi_1^{\times}(x) \underset{5.91}{=} x$
6. If $x \in \prod_{i \in I} A_i$ [the general not finite case] then we use the notation referenced in 2.84 where x_i means $x(i)$
7. Let I, A be sets then $A^I = \prod_{i \in I} C_i$ where $\{C_i\}_{i \in I}$ is defined by $C_i = A$ (this is actually a theorem see 2.82)
8. If $n \in \mathbb{N}$ then and A a set then $A^n = \prod_{i \in \{1,\ldots n\}} C_i$ [$\underset{(1)}{=} \bigotimes_{i \in \{1,\ldots,n\}} C_i$] where $\{C_i\}_{i \in \{1,\ldots,n\}}$ is define by $C_i = A$

Using the above theorem we can restate the basic theorems about product of sets

1. Let I be a set and $\{A_i\}_{i \in I}$ a family of **non empty** sets then $\forall i \in I$ π_i: $\prod_{i \in I} A_i \to A_i$ is a surjection (see 2.89)
2. Let $n \in \mathbb{N}$, $\{A_i\}_{i \in \{1,\ldots,n\}}$ a finite family of non empty sets then $\forall i \in \{1, \ldots, n\}$ π_i: $\prod_{i \in \{1,\ldots,n\}} A_i \to A_i$ is a surjection (see 5.92 (2))
3. Let $n \in \mathbb{N}$, $\{A_i\}_{i \in \{1,\ldots,n+1\}}$ a finite family of sets then $\prod_{i \in \{1,\ldots,n+1\}} A_i = \left(\prod_{i \in \{1,\ldots,n\}} A_i\right) \times A_{n+1}$ and $\forall (x_1, \ldots, x_n) \in \prod_{i \in \{1,\ldots,n\}} A_i$ and $x \in A_{n+1}$ we have that $((x_1, \ldots, x_n), x) = (x_1, \ldots, x_{n+1})$ (see 5.81)
4. Let $n, k \in \mathbb{N}$ with $1 \leqslant k < n$ and $\{A_i\}_{i \in \{1,\ldots,n\}}$ then there exists a bijection β: $\prod_{i \in \{1,\ldots,n\}} A_i \to \left(\prod_{i \in \{1,\ldots,k\}} A_i\right) \times \left(\prod_{i \in \{1,\ldots,n-k\}} A_{i+k}\right)$ such that $\beta((x_1, \ldots, x_n)) = ((x_1, \ldots, x_k), (x_{k+1}, \ldots, x_n))$
5. Let I be a set and $\{A_i\}_{i \in I}$, $\{B_i\}_{i \in I}$ two families of sets such that $\forall x \in \prod_{i \in I} A_i$ we have that $\forall i \in I$ $x_i \in B_i$ then $\prod_{i \in I} A_i \subseteq \prod_{i \in I} B_i$ (see 2.88)
6. Let $n \in \mathbb{N}$ and $\{A_i\}_{i \in \{1,\ldots,n\}}$, $\{B_i\}_{i \in \{1,\ldots,n\}}$ two finite families of sets such that $\forall x \in \prod_{i \in \{1,\ldots,n\}} A_i$ we have that $\forall i \in \{1, \ldots, n\}$ $x_i \in B_i$ then $\prod_{i \in \{1,\ldots,n\}} A_i \subseteq \prod_{i \in I} B_i$ (see 5.92 (3))
7. Let I be a set and $\{A_i\}_{i \in I}$ a family of sets then $\prod_{i \in I} A_i = \emptyset \Leftrightarrow \exists i \in I$ such that $A_i = \emptyset$ (see 2.86)
8. Let $n \in \mathbb{N}$ and $\{A_i\}_{i \in I}$ a finite family of sets then $\prod_{i \in \{1,\ldots,n\}} A_i = \emptyset \Leftrightarrow \exists i \in \{1, \ldots, n\}$ such that $A_i = \emptyset$ (see 5.89)
9. Let I be a set, $\{A_i\}_{i \in I}$, $\{B_i\}_{i \in I}$ two families of sets then we have
   a. If $\forall i \in I$ we have $A_i \subseteq B_i$ then $\prod_{i \in I} A_i \subseteq \prod_{i \in I} B_i$ (see 2.85)
   b. If $\forall i \in I$ we have $A_i \neq \emptyset$ then $\prod_{i \in I} A_i \subseteq \prod_{i \in I} B_i \Leftrightarrow \forall i \in I \vDash A_i \subseteq B_i$ (see 2.91)

10. Let $n\in\mathbb{N}$, $\{A_i\}_{i\in\{1,\ldots,n\}}$, $\{B_i\}_{i\in\{1,\ldots,n\}}$ two finite families of sets then we have

    a. If $\forall i\in\{1,\ldots,n\}$ we have $A_i\subseteq B_i$ then $\prod_{i\in\{1,\ldots,n\}} A_i\subseteq\prod_{i\in\{1,\ldots,n\}} B_i$ (see 5.83)

    b. If $\forall i\in\{1,\ldots,n\}$ we have $A_i\neq\emptyset$ then $\prod_{i\in\{1,\ldots,n\}} A_i\subseteq\prod_{i\in I} B_i\Leftrightarrow\forall i\in\{1,\ldots,n\}\vDash A_i\subseteq B_i$ (see 5.94)

11. Let I,J be sets and $\{\{A_{i,j}\}_{j\in J}\}_{i\in I}$ a family of families of sets then $\bigcap_{i\in I}(\prod_{j\in J}A_{i,j})=\prod_{j\in J}(\bigcap_{i\in I}A_{i,j})$ (see 2.96)

12. Let I be a set, $n\in\mathbb{N}$ and $\{\{A_{i,j}\}_{j\in\{1,\ldots,n\}}\}_{i\in I}$ a family of finite families of sets then $\bigcap_{i\in I}\left(\prod_{j\in\{1,\ldots,n\}}A_{i,j}\right)=\prod_{j\in\{1,\ldots,n\}}(\bigcap_{i\in I}A_{i,j})$ (see 5.97)

13. If I is a set and $\{A_i\}_{i\in I}$, $\{B_i\}_{i\in I}$ families of sets then $(\prod_{i\in I}A_i)\bigcup(\prod_{i\in I}B_i)\subseteq\prod_{i\in I}(A_i\bigcup B_i)$ (see 2.97)

14. If $n\in\mathbb{N}$ and $\{A_i\}_{i\in\{1,\ldots,n\}}$, $\{B_i\}_{i\in\{1,\ldots,n\}}$ finite families of sets then $\left(\prod_{i\in\{1,\ldots,n\}}A_i\right)\bigcup\left(\prod_{i\in\{1,\ldots,n\}}B_i\right)\subseteq\prod_{i\in\{1,\ldots,n\}}(A_i\bigcup B_i)$ (see 5.98)

We have the following corollary for (8) and (9) above

Corollary 5.103. *Equality for a product of sets*

1. *If I is a set and $\{A_i\}_{i\in I},\{B_i\}_{i\in I}$ be two family of sets such that $\forall i\in I$ $A_i\neq\emptyset$ then $\prod_{i\in I}A_i=\prod_{i\in I}B_i\Leftrightarrow\forall i\in I$ we have $A_i=B_i$*

2. *If $n\in\mathbb{N}$ and $\{A_i\}_{i\in\{1,\ldots,n\}},\{B_i\}_{i\in\{1,\ldots,n\}}$ be two family of sets such that $\forall i\in\{1,\ldots,n\}$ $A_i\neq\emptyset$ then $\prod_{i\in\{1,\ldots,n\}}A_i=\prod_{i\in\{1,\ldots,n\}}B_i\Leftrightarrow\forall i\in I$ we have $A_i=B_i$*

Proof.

1.

   ⇒. As $\forall i\in I$ we have $A_i\neq\emptyset$ we have by 5.102 (6) that $\prod_{i\in I}A_i\neq\emptyset$ hence as $\prod_{i\in I}A_i=\prod_{i\in I}B_i$ we have by 5.102 (6) that $\forall i\in I$ $B_i\neq\emptyset$.From $\prod_{i\in I}A_i=\prod_{i\in I}B_i$ we have $\prod_{i\in I}A_i\subseteq\prod_{i\in I}B_i\underset{5.102\ (8)}{\Rightarrow}\forall i\in I\quad A_i\subseteq B_i$ and $\prod_{i\in I}B_i\subseteq\prod_{i\in I}A_i\underset{5.102\ (8)}{\Rightarrow}\forall i\in I B_i\subseteq A_i$. So we conclude that $\forall i\in I$ $A_i=B_i$.

   ⇐. This is trivial

2.

   ⇒. As $\forall i\in\{1,\ldots,n\}$ we have $A_i\neq\emptyset$ we have by 5.102 (6) that $\prod_{i\in\{1,\ldots,n\}}A_i\neq\emptyset$ hence as $\prod_{i\in\{1,\ldots,n\}}A_i=\prod_{i\in\{1,\ldots,n\}}B_i$ we have by 5.102 (6) that $\forall i\in\{1,\ldots,n\}$ $B_i\neq\emptyset$.From $\prod_{i\in\{1,\ldots,n\}}A_i=\prod_{i\in\{1,\ldots,n\}}B_i$ we have $\prod_{i\in\{1,\ldots,n\}}A_i\subseteq\prod_{i\in\{1,\ldots,n\}}B_i\underset{5.102\ (8)}{\Rightarrow}\forall i\in\{1,\ldots,n\}$ $A_i\subseteq B_i$ and $\prod_{i\in\{1,\ldots,n\}}B_i\subseteq\prod_{i\in\{1,\ldots,n\}}A_i\underset{5.102\ (8)}{\Rightarrow}\forall i\in\{1,\ldots,n\}$ $B_i\subseteq A_i$. So we conclude that $\forall i\in\{1,\ldots,n\}$ $A_i=B_i$.

   ⇐. This is trivial □

Theorem 5.104. *Let $n, m \in \mathbb{N}$ and $\{A_i\}_{i \in \{1,\ldots,m\}}$ a family of* ***non empty*** *sets such that*

$$\forall i \in \{1,\ldots,m\} \text{ there exists a } \{B_{i,j}\}_{j \in \{1,\ldots,n\}} \text{ such that } A_i = \prod_{j \in \{1,\ldots,n\}} B_{i,j}$$

$$\exists \{C_i\}_{i \in \{1,\ldots,n\}} \text{ such that } \bigcup_{i \in \{1,\ldots,m\}} A_i = \prod_{i \in \{1,\ldots,n\}} C_i$$

then we have

1. $\bigcup_{i \in \{1,\ldots,m\}} A_i = \prod_{j \in \{1,\ldots,n\}} \left(\bigcup_{i \in \{1,\ldots,m\}} B_{i,j}\right)$
2. $\forall i \in \{1,\ldots,n\}$ $C_i = \bigcup_{j \in \{1,\ldots,m\}} B_{j,i}$

Proof. First

$$\begin{aligned} x = (x_1,\ldots,x_n) \in \bigcup_{i \in \{1,\ldots,m\}} A_i &\Rightarrow \exists i \in \{1,\ldots,m\} \vdash (x_1,\ldots,x_n) \in A_i \\ &\Rightarrow \forall j \in \{1,\ldots,n\} \vdash x_j \in B_{i,j} \subseteq \bigcup_{k \in \{1,\ldots,m\}} B_{k,j} \\ &\Rightarrow (x_1,\ldots,x_n) \in \prod_{j \in \{1,\ldots,n\}} \left(\bigcup_{k \in \{1,\ldots,m\}} B_{k,j}\right) \end{aligned}$$

proving

$$\bigcup_{i \in \{1,\ldots,m\}} A_i \subseteq \prod_{j \in \{1,\ldots,n\}} \left(\bigcup_{k \in \{1,\ldots,m\}} B_{k,j}\right) \tag{5.12}$$

Second as $\forall i \in \{1,\ldots,m\}$ $\emptyset \neq A_i$ if follows from 5.102 (7) that

$$\forall j \in \{1,\ldots,n\} \vDash B_{i,j} \neq \emptyset \tag{5.13}$$

Given $i \in \{1,\ldots,m\}$ take $j \in \{1,\ldots,n\}$ and $x_j \in B_{i,j}$ then by the above there exists $\forall k \in \{1,\ldots.,n\} \setminus \{j\}$ a $x_k \in B_{i,k}$ so that $(x_1,\ldots,x_k) \in \prod_{l \in \{1,\ldots,n\}} B_{i,j} = A_i \subseteq \bigcup_{i \in \{1,\ldots,m\}} A_i = \prod_{l \in \{1,\ldots,n\}} C_l$ proving that $x_j \in C_j$ hence

$$\forall j \in \{1,\ldots,n\} \vDash x_j \in B_{i,j} \Rightarrow x \in C_j \tag{5.14}$$

Next

$$\begin{aligned} x = (x_1,\ldots,x_n) \in \prod_{j \in \{1,\ldots,n\}} \left(\bigcup_{k \in \{1,\ldots,m\}} B_{k,j}\right) &\Rightarrow \forall j \in \{1,\ldots,n\} \vDash x_j \in \bigcup_{k \in \{1,\ldots,m\}} B_{k,j} \\ &\Rightarrow \exists k \in \{1,\ldots,n\} \vdash x_j \in B_{k,j} \\ &\underset{5.14}{\Rightarrow} x_j \in C_j \\ &\Rightarrow (x_1,\ldots,x_n) \in \prod_{j \in \{1,\ldots,n\}} C_j = \bigcup_{i \in \{1,\ldots,m\}} A_i \end{aligned}$$

proving that $\prod_{j\in\{1,\ldots,n\}}\left(\bigcup_{k\in\{1,\ldots,m\}} B_{k,j}\right)\subseteq\bigcup_{i\in\{1,\ldots,m\}} A_i$ and using 5.12 this gives

$$\prod_{j\in\{1,\ldots,n\}}\left(\bigcup_{k\in\{1,\ldots,m\}} B_{k,j}\right)=\bigcup_{i\in\{1,\ldots,m\}} A_i$$

proving (1) of the lemma. Finally (2) follows from 5.103 $\square$

Chapter 6

The Integer numbers

6.1 Definition and arithmetic's

The problem with the natural numbers is that we have no inverse number for every natural number, meaning that $\mathbb{N}_0$ can not be a group. In other words given $n \in \mathbb{N}_0$ with $n \neq 0$ there does not exists a n' such that $n + n' = 0$. This is solved by introducing the whole numbers.

Definition 6.1. *We define the relation* $\sim$ *on* $\mathbb{N}_0 \times \mathbb{N}_0$ *by* $\sim = \{((n,m),(n',m')) \in (\mathbb{N}_0 \times \mathbb{N}_0) \times (\mathbb{N}_0 \times \mathbb{N}_0) | n+m' = m+n'\}$

Theorem 6.2. $\sim$ *is a equivalence relation on* $\mathbb{N}_0$

Proof.

1. **(reflectivity)** If $(n,m) \in \mathbb{N}_0 \times \mathbb{N}_0$ then $n+m \underset{4.33}{=} m+n \Rightarrow (n,m) \sim (n,m)$
2. **(Symmetry)** If $(n,m) \sim (n',m') \Rightarrow n+m' = m+n' \underset{4.33}{\Rightarrow} m'+n = n'+m \Rightarrow n'+m = m'+n \Rightarrow (n',m') \sim (n,m)$
3. **(Transitivity)** If $(n,m) \sim (n',m')$ and $(n',m') \sim (n'',m'') \Rightarrow n+m' = m+n'$ and $n'+m'' = m'+n'' \Rightarrow (n+m')+(n'+m'') = (m+n')+(m'+n'') \underset{4.33,4.32}{\Rightarrow} (n+m'')+(m'+n') = (m+n'')+(m'+n') \underset{4.44}{\Rightarrow} n+m'' = m+n'' \Rightarrow (n,m) \sim (n'',m'')$ □

Definition 6.3. *We define the set of integer numbers* $\mathbb{Z}$ *by* $\mathbb{Z} = \mathbb{N}_0/\sim$ *(see 2.119)*

So we have that $\mathbb{Z} = \{\sim[(n,m)] | (n,m) \in \mathbb{N}_0 \times \mathbb{N}_0\}$

Theorem 6.4. *If* $\sim[(n,m)] \in \mathbb{Z}$ *then if* $k \in \mathbb{N}_0$ *we have* $\sim[(n,m)] = \sim[(n+k,m+k)]$

Proof. $n+(m+k) = (n+m)+k = (m+n)+k = m+(n+k) \Rightarrow (n,m) \sim (n+k,m+k) \Rightarrow \sim[(n,m)] = \sim[(n+k,m+k)]$ □

Theorem 6.5. $+ \colon \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ *defined by* $(\sim[(n,m)], \sim[r,s]) \to \sim[(n,m)] + \sim[(r,r)] = \sim[(n+r,m+s)]$ *is a function. Note that we have here two additions + one defined in* $\mathbb{N}_0$ *the other in* $\mathbb{Z}$.

Proof. We must prove that if $\sim[(n,m)]=\sim[(n',m')]$ and $\sim[(r,s)]=\sim[(r',s')]$ then $\sim[(n+r),(m+s)]=\sim[(n'+r',m'+s')]$. Now from $\sim[(n,m)]=\sim[(n',m')]$ and $\sim[(r,s)]=\sim[(r',s')]$ we have by 2.116 that $(n,m)\sim(n',m')$ and $(r,s)\sim(r',s')\Rightarrow n+m'=m+n'$ and $r+s'=s+r'$. So $(n+r)+(m'+s')=(n+m')+(r+s')=(m+n')+(s+r')=(m+s)+(n'+r')\Rightarrow(n+r,m+s)\sim(n'+r',m'+s')\Rightarrow\sim[(n+r,m+s)]=\sim[(n'+r',m'+s')]$ □

Theorem 6.6. **($\langle\mathbb{Z},+\rangle$ is a group)** *We have the following properties for* $+:\mathbb{Z}\times\mathbb{Z}\rightarrow\mathbb{Z}$

1. **(Associativity)** $\forall n,m,k\in\mathbb{Z}$ *we have* $(x+y)+z=x+(y+z)$
2. **(Neutral element)** $\forall n\in\mathbb{Z}$ *there exists a* $0\in\mathbb{Z}$ *with* $n+0=0+n=n$ *(here* $0=\sim[(1,1)]$*, note that* 0 *here is not the same as the* 0 *in* $\mathbb{N}_0$*, although the* $1\in\mathbb{N}_0$*)*
3. **(Commutativity)** $\forall n,m\in\mathbb{Z}$ *then* $n+m=m+n$
4. **(Inverse Element)** $\forall n\in\mathbb{Z}$ *there exists a* ***unique*** *element* $-n$ *such that* $(-n)+n=0=n+(-n)$*. If* $n=\sim[(n_1,n_2)]\Rightarrow-n=\sim[(n_2,n_1)]$ *(note as inverse elements are always unique we don't have to prove that* $-n$ *is defined independent of its representation).*

Proof.

1. If $n=\sim[(n_1,n_2)]$, $m=\sim[(m_1,m_2)]$ and $k=\sim[(k_1,k_2)]$ then $(n+m)+k=\sim[(n_1+m_1,n_2+m_2)]+k=\sim[((n_1+m_1)+k_1,(n_2+m_2)+k_2)]\underset{4.32}{=}\sim[(n_1+(m_1+k_1),n_2+(m_2+k_2))]=\sim[(n_1,n_2)]+\sim[(m_1+k_1,m_2+k_2)]=n+([m_1,m_2]+\sim[(k_1,k_2)])=n+(m+k)$
2. If $n=\sim[(n_1,n_2)]$ then $\sim[(1,1)]+\sim[(n_1,n_1)]=\sim[(1+n_1,1+n_1)]$ now $(1+n_1)+n_2\underset{4.32,4.33}{=}(1+n_2)+n_1\Rightarrow(1+n_1,1+n_2)\sim(n_1,n_2)\Rightarrow\sim[(1,1)]+\sim[(n_1,n_1)]=\sim[(n_1,n_2)]$. Also $\sim[(n_1,n_2)]+\sim[(1,1)]=\sim[(n_1+1,n_2+1)]\underset{4.33}{=}\sim[(1+n_1,1+n_2)]\underset{\text{already proved}}{=}\sim[(n_1,n_2)]$
3. If $n=\sim[(n_1,n_2)]$ and $m=\sim[(m_1,m_2)]$ then $n+m=\sim[(n_1+m_1,n_2+m_2)]\underset{4.33}{=}\sim[(m_1+n_1,m_2+n_2)]=m+n$
4. If $n=\sim[(n_1,n_2)]$ define then $-n=\sim[(n_2,n_1)]$ then $n+(-n)=\sim[(n_1+n_2,n_2+n_1)]\underset{4.33}{=}\sim[(n_1+n_2,n_1+n_2)]$ now $(n_1+n_2)+1=(n_1+n_2)+1\Rightarrow(n_1+n_2,n_1+n_2)=\sim[(1,1)]\Rightarrow n+(-n)=0$ and by (3) we have $(-n)+n=0$ □

Theorem 6.7. *If* $n\in\mathbb{N}_0$ *then* $0=\sim[(n,n)]$

Proof. $0=\sim[(1,1)]$ and $n+1=n+1\Rightarrow(1,1)\sim(n,n)\Rightarrow\sim[(n,n)]=\sim[(1,1)]$ □

Theorem 6.8. $\cdot:\mathbb{Z}\times\mathbb{Z}\rightarrow\mathbb{Z}$ *defined by* $(\sim[(m,n)],\sim[(k,r)])\rightarrow\sim[(m,n)]\cdot\sim[(k,r)]=\sim[(m\cdot k+n\cdot r,n\cdot k+m\cdot r)]$ *is a function. Note that we have two multiplications* $\cdot$ *here one defined in* $\mathbb{N}_0$ *the other in* $\mathbb{Z}$*.*

Proof. If $(m,n)\sim(m',n')$ and $(k,r)\sim(k',r')$ then we have $m+n'=n+m'\wedge k+r'=r+k'$ so we have then

$$\begin{array}{rcl}
(m+n')\cdot[(r+k')+(r+k')] & = & (m+n')\cdot(r+k')+(m+n')\cdot(r+k') \\
 & = & [(m+n')+(m+n')]\cdot(r+k') \\
 & \Leftrightarrow & \\
(m+n')\cdot[(r+k')+(k+r')] & = & [(m+n')+(n+m')]\cdot(r+k') \\
 & \Leftrightarrow & \\
(m+n')\cdot[k+k'+r'+r] & = & [n+m+m'+n']\cdot(r+k') \\
 & \Leftrightarrow & \\
(m+n')\cdot k+(m+n')\cdot k'+(m'+n)\cdot r'+(m'+n)\cdot r & = & n\cdot(r'+k)+m\cdot(r+k')+m'\cdot(r+k')+n'\cdot(r'+k) \\
m\cdot k+n\cdot r+n'\cdot k'+m'\cdot r'+(n'\cdot k+m'\cdot r+n\cdot r'+m\cdot k') & = & n\cdot k+m\cdot r+m'\cdot k'+n'\cdot r'+(n'\cdot k+m'\cdot r+n\cdot r'+m\cdot k') \\
 & \Leftrightarrow & \\
m\cdot k+n\cdot r+n'\cdot k'+m'\cdot r' & = & n\cdot k+m\cdot r+m'\cdot k'+n'\cdot r' \\
 & \Leftrightarrow & \\
(m\cdot k+n\cdot r)+(n'\cdot k'+m'\cdot r') & = & (n\cdot k+m\cdot r)+(m'\cdot k'+n'\cdot r') \\
 & \Leftrightarrow & \\
(m\cdot k+n\cdot r, n\cdot k+m\cdot r) & \sim & (m'\cdot k'+n'\cdot r', n'\cdot k'+m'\cdot r') \\
\sim[(m\cdot k+n\cdot r, n\cdot k+m\cdot r)] & = & \sim[(m'\cdot k'+n'\cdot r', n'\cdot k'+m'\cdot r')]
\end{array}$$

proving that our definition is independent of the representation and that $\cdot$ is a function.

□

Theorem 6.9. **($\langle\mathbb{Z},.\rangle$ is a abelian semi-group)** *If $n,m,k\in\mathbb{Z}$ then we have*

1. **(Commutativity)** $n\cdot m=m\cdot n$
2. **(Associativity)** $n\cdot(m\cdot k)=(n\cdot m)\cdot k$
3. **(Neutral Element)** *There exist a $1=\sim[(2,1)]$ such that $n\cdot 1=1\cdot n=n$ (note that we have two 1 here a $1\in\mathbb{Z}$ and a $1\in\mathbb{N}_0$)*
4. **(Distributivity)** $n\cdot(m+k)=n\cdot m+n\cdot k$
5. **(There does not exists a zero divisor)** $n\cdot m=0\Rightarrow n=0\vee m=0$
6. $(-1)\cdot(-1)=1$

Note that the symbol 1 is here used as the unit of $\mathbb{N}_0$ and $\mathbb{Z}$, of course these units are different objects, but context will always tell if $1\in\mathbb{N}_0$ or $1=\sim[(2,1)]\in\mathbb{Z}$.

Proof. Let $n=\sim[(n_1,n_2)]$, $m=\sim[(m_1,m_2)]$ and $k=\sim[(k_1,k_2)]$ then we have

1. $n\cdot m=\sim[(n_1,n_2)]\cdot\sim[(m_1,m_2)]=\sim[(n_1\cdot m_1+n_2\cdot m_2, n_2\cdot m_1+n_1\cdot m_2)]\underset{4.41}{=}\sim[(m_1\cdot n_1+m_2\cdot n_2, m_1\cdot n_2+m_2\cdot n_1)]=$

2. $n\cdot(m\cdot k)=n\cdot\sim[(m_1\cdot k_1+m_2\cdot k_2,m_2\cdot k_1+m_1\cdot k_2)]=\sim[(n_1\cdot(m_1\cdot k_1+m_2\cdot k_2)+n_2\cdot(m_2\cdot k_1+m_1\cdot k_2),n_2\cdot(m_1\cdot k_1+m_2\cdot k_2)+n_1\cdot(m_2\cdot k_1+m_1\cdot k_2))]=\sim[((n_1\cdot m_1)\cdot k_1+(n_1\cdot m_2)\cdot k_2+(n_2\cdot m_2)\cdot k_1+(n_2\cdot m_1)\cdot k_2,(n_2\cdot m_1)\cdot k_1)+(n_2\cdot m_2)\cdot k_2+(n_1\cdot m_2)\cdot k_1+(n_1\cdot m_1)\cdot k_2]=\sim[((n_1\cdot m_1+n_2\cdot m_2)\cdot k_1+(n_1\cdot m_2+n_2\cdot m_1)\cdot k_2,(n_2\cdot m_1+n_1\cdot m_2)\cdot k_1+(n_2\cdot m_2+n_1\cdot m_1)\cdot k_2)]=\sim[(n_1\cdot m_1+n_2\cdot m_2,n_1\cdot m_2+n_2\cdot m_1)]\cdot\sim[(k_1,k_2)]=(\sim[(n_1,n_2)]\cdot\sim[(m_1,m_2)])\cdot\sim[(k_1,k_2)]=(n\cdot m)\cdot k$

3. $1\cdot n=\sim[2,1]\cdot\sim[(n_1,n_2)]=\sim[(2\cdot n_1+1\cdot n_2,1\cdot n_1+2\cdot n_2)]=\sim[((1+1)\cdot n_1+n_2,n_1+(1+1)\cdot n_2)]=\sim[(n_1+n_1+n_2,n_1+n_2+n_2)]=\sim[(n_1,n_2)]=n$ as $n_1+n_1+n_2+n_2=n_1+n_2+n_2+n_1$. The fact that $n\cdot 1=n$ follows from (1).

4. $n\cdot(m+k)=n\cdot\sim[(m_1+k_1,m_2+k_2)]=\sim[(n_1\cdot(m_1+k_1)+n_2\cdot(m_2+k_2),n_2\cdot(m_1+k_1)+n_1\cdot(m_2+k_2))]=\sim[(n_1\cdot m_1+n_1\cdot k_1+n_2\cdot m_2+n_2\cdot k_2,n_2\cdot m_1+n_2\cdot k_1+n_1\cdot m_2+n_1\cdot k_2)]=\sim[((n_1\cdot m_1+n_2\cdot m_2+n_1\cdot k_1+n_2\cdot k_2,n_1\cdot m_2+n_2\cdot m_1+n_2\cdot k_1+n_1\cdot k_2))]=\sim[(n_1\cdot m_1+n_2\cdot m_2,n_1\cdot m_2+n_2\cdot m_1)]+\sim[(n_1\cdot k_1+n_2\cdot k_2,n_2\cdot k_1+n_1\cdot k_2)]=\sim[(n_1,n_2)]\cdot\sim[(m_1,m_2)]+\sim[(n_1,n_2)]\cdot\sim[(k_1,k_2)]=n\cdot m+n\cdot k$

5. If $n\cdot m=0$ then $\sim[(n_1\cdot m_1+n_2\cdot m_2,n_2\cdot m_1+n_1\cdot m_2)]=\sim[(1,1)]\Rightarrow$

$$\begin{array}{rcl} n_1\cdot m_1+n_2\cdot m_2+1 & = & n_2\cdot m_1+n_1\cdot m_2+1\\ & \underset{4.44}{\Rightarrow} & \\ n_1\cdot m_1+n_2\cdot m_2 & = & n_2\cdot m_1+n_1\cdot m_2 \end{array}$$

Suppose now that $\sim[(n_1,n_2)]\neq\sim[(1,1)]\Rightarrow n_1+1\neq n_2+1$ then if $n_1=n_2\Rightarrow n_1+1=n_2+1$ a contradiction so we have $n_1\neq n_2$. So the following cases are possible

a. $n_1<n_2\underset{4.61}{\Rightarrow}\exists l\in\mathbb{N}\Rightarrow n_1+l=n_2\Rightarrow n_1\cdot m_1+(n_1+l)\cdot m_2=(n_1+l)\cdot m_1+n_1\cdot m_2\Rightarrow l\cdot m_2+(n_1\cdot m_1+n_1\cdot m_2)=l\cdot m_1+(n_1\cdot m_1+n_1\cdot m_2)\underset{4.44}{\Rightarrow}l\cdot m_1=l\cdot m_2\underset{4.72}{\Rightarrow}m_1=m_2\Rightarrow m_1+1=m_2+1\Rightarrow\sim[(m_1,m_2)]=\sim[(1,1)]=0$

b. $n_2<n_1\underset{4.61}{\Rightarrow}\exists l\in\mathbb{N}\Rightarrow n_2+l=n_1\Rightarrow(n_2+l)\cdot m_1+n_2\cdot m_2=n_2\cdot m_1+(n_2+l)\cdot m_2\Rightarrow l\cdot m_1+(n_2\cdot m_1+n_2\cdot m_2)=l\cdot m_2+(n_2\cdot m_1+n_2\cdot m_2)\underset{4.44}{\Rightarrow}l\cdot m_2=l\cdot m_1\underset{4.72}{\Rightarrow}m_2=m_1\Rightarrow m_2+1=m_1+1\Rightarrow\sim[(m_1,m_2)]=\sim[(1,1)]=0$

6. $(-1)\cdot(-1)=\sim[(1,2)]\cdot\sim[(1,2)]=\sim[(1\cdot 1+2\cdot 2,2\cdot 1+1\cdot 2)]=\sim[(5,4)]=\sim[(2,1)]=1$ ☐

From 6.6 and 6.9 we have that

Theorem 6.10. *$\langle\mathbb{Z},+,\cdot\rangle$ is a **ring** even more $\langle\mathbb{Z},+,\cdot\rangle$ forms a **integral domain***

Notation 6.11. *Whenever we write $a-b$ we actually mean $a+(-b)$ (where $-b$ is the inverse of b in the group $\langle\mathbb{Z},+\rangle$.*

Theorem 6.12. (Absorbing element) *If $n\in\mathbb{Z}$ then $0\cdot n=0=n\cdot 0$*

Proof. If $n = \sim[(n_1, n_2)]$ then we have

$$\begin{aligned} 0 \cdot n &= \sim[(1,1)] \cdot \sim[(n_1, n_2)] \\ &= \sim[(1 \cdot n_1 + 1 \cdot n_2, 1 \cdot n_1 + 1 \cdot n_2)] \\ &= \sim[(n_1 + n_2, n_1 + n_2)] \\ &\underset{6.4}{=} \sim[(n_1, n_1)] = \sim[(0 + n_1, 0 + n_1)] \\ &\underset{6.4}{=} \sim[(0,0)] \\ &\underset{6.4}{=} \sim[(0+1, 0+1)] \\ &\underset{6.4}{=} \sim[(1,1)] \\ &= 0 \end{aligned}$$

The remaining is proved by commutativity of $\cdot$

□

Lemma 6.13. *If $n \in \mathbb{Z}$ then $-n = (-1) \cdot n$*

Proof. First $1 = \sim[(2,1)] \Rightarrow -1 = \sim[(1,2)] \underset{6.4}{=} \sim[(0,1)]$. Now if $n \in \mathbb{Z} \Rightarrow n = \sim[(n_1, n_2)]$ then we have

$$\begin{aligned} (-1) \cdot n &= \sim[(0 \cdot n_1 + 1 \cdot n_2, 0 \cdot n_2 + 1 \cdot n_1)] \\ &= \sim[(n_2, n_1)] \\ &= -n \end{aligned}$$

□

Theorem 6.14. *If $n, m \in \mathbb{Z}$ then $-(n \cdot m) = (-n) \cdot m = n \cdot (-m)$*

Proof. $-(n \cdot m) = (-1) \cdot (n \cdot m) = ((-1) \cdot n) \cdot m \underset{6.13}{=} (-n) \cdot m$, also $-(n \cdot m) = -(m \cdot n) = (-m) \cdot n = n \cdot (-m)$ □

Theorem 6.15. *If $n, k, r \in \mathbb{Z}$ and $r \neq 0$ then we have $n \cdot r = k \cdot r \Rightarrow n = k$*

Proof. From $n \cdot r = k \cdot r$ we get $n \cdot r + (-(k \cdot r)) = k \cdot r + (-(k \cdot r)) = 0 \underset{6.14}{\Rightarrow} n \cdot r + (-k) \cdot r = 0 \underset{6.9}{\Rightarrow} (n + (-k)) \cdot r = 0 \underset{6.9 \text{ and } r \neq 0}{\Rightarrow} n + (-k) = 0 \Rightarrow (n + (-k)) + k = k \Rightarrow n + ((-k) + k) = k \Rightarrow n + 0 = k \Rightarrow n = k$ □

6.1.1 Power in $\mathbb{Z}$

Definition 6.16. *As $\langle \mathbb{Z}, \cdot \rangle$ is a abelian semi-group we have by 4.22 that given a $a \in \mathbb{Z}$ and $n \in \mathbb{N}_0$ that there exists a a^n such that*

$$\begin{aligned} a^0 &= 1 \\ a^{n+1} &= a^n \cdot a \underset{\text{abelian}}{=} a \cdot a^n \end{aligned}$$

Theorem 6.17. *If $n, n' \in \mathbb{N}_0$ and $a \in \mathbb{Z}$ then $a^{n'+n} = a^{n'} \cdot a^n$*

Proof. We prove this by induction on n. So let $X = \{n \in \mathbb{N}_0 | a^{n'+n} = a^{n'} \cdot a^n\}$ then we have

1. If $n = 0$ then $a^{n'+n} = a^{n'+0} = a^{n'} = a^{n'} \cdot 1 = a^{n'} \cdot a^0 \Rightarrow 0 \in X$

2. If $n \in X$ then $a^{n'+(n+1)} = a^{(n'+n)+1} = a^{(n'+n)} \cdot a \underset{n \in X}{=} (a^{n'} \cdot a^n) \cdot a = a^{n'} \cdot (a^n \cdot a) = a^{n'} \cdot a^{n+1}$ and thus $n+1 \in X$

Using mathematical induction (see 4.10) we have $X = \mathbb{N}_0$ proving the theorem $\square$

Theorem 6.18. *If $n \in \mathbb{N}_0$ and $a, b \in \mathbb{Z}$ then we have $(a \cdot b)^n = a^n \cdot b^n$*

Proof. We prove this by induction so take $\mathcal{S} = \{n \in \mathbb{N}_0 | (a \cdot b)^n = a^n \cdot b^n\}$ then we have

$\boldsymbol{n = 0.}$ then $(a \cdot b)^0 = 1 = 1 \cdot 1 = a^0 \cdot b^0$

$\boldsymbol{n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}.}$ then $(a \cdot b)^{n+1} = (a \cdot b)^n \cdot (a \cdot b) = (a^n \cdot b^n) \cdot (a^n \cdot a) \cdot (b^n \cdot b) = a^{n+1} \cdot b^{n+1}$ $\square$

Theorem 6.19. *In $\mathbb{Z}$ we have*

$$\begin{aligned} 0^n &= 0 \,(\text{if } n \neq 0) \\ 1^n &= 1 \\ (-1)^n &= -1 \text{ or } 1 \\ (-1)^{2 \cdot n} &= 1 \\ (-1)^{2 \cdot n+1} &= -1 \end{aligned}$$

Proof.

1. If $n \neq 0 \Rightarrow \exists m \in \mathbb{N}_0 \vdash n = m+1$ then $0^n = 0^{(m+1)} = 0^m \cdot 0 = 0$
2. $1^n = 1$ is proved by induction on n, let $X = \{n \in \mathbb{N}_0 | 1^n = 1\}$ then
    a. $1^0 = 1 \Rightarrow 0 \in X$
    b. If $n \in X \Rightarrow 1^{n+1} = 1^n \cdot 1 \underset{n \in X}{=} 1 \cdot 1 = 1 \Rightarrow n+1 \in X$

    so $X = \mathbb{N}_0$
3. $(-1)^n = \pm 1$ is proved by induction on n, let $X = \{n \in \mathbb{N}_0 | (-1)^n = -1 \text{ or } 1\}$ then
    a. $(-1)^0 = 1 \Rightarrow 0 \in X$
    b. If $n \in X$ then $(-1)^{n+1} = (-1)^n \cdot (-1) \underset{n \in X}{=} (-1) \cdot (-1) \vee 1 \cdot (-1) = 1 \vee -1 \Rightarrow n+1 \in X$

    so $X = \mathbb{N}_0$
4. $(-1)^{2 \cdot n} = (-1)^{(1+1) \cdot n} = (-1)^{n+n} = (-1)^n \cdot (-1)^n \underset{(3)}{=} (-1) \cdot (-1) \text{ or } 1 \cdot 1 = 1$
5. $(-1)^{2 \cdot n+1} = (-1)^{2 \cdot n} \cdot (-1) \underset{(4)}{=} 1 \cdot (-1) = -1$ $\square$

6.2 Order relation of the Integers

Definition 6.20. $\mathbb{N}_{0_{\mathbb{Z}}} = \{\sim[(s(n), 1)] | n \in \mathbb{N}_0\} \subseteq \mathbb{Z}$, *$\mathbb{N}_{0_{\mathbb{Z}}}$ is called the collection of positive whole numbers.*

Note 6.21. $0 = \sim[(1,1)] \in \mathbb{N}_{0_{\mathbb{Z}}}, 1 = \sim[(2,1)] \in \mathbb{N}_{0_{\mathbb{Z}}}$

Theorem 6.22. *We have the following*

1. $\langle \mathbb{N}_{0_{\mathbb{Z}}}, + \rangle$ *is a sub-semi-group of* $\langle \mathbb{Z}, + \rangle$
2. $\langle \mathbb{N}_{0_{\mathbb{Z}}}, . \rangle$ *is a sub-semi-group of* $\langle \mathbb{Z}, . \rangle$
3. $i_{\mathbb{N}_0}: \mathbb{N}_0 \rightarrow \mathbb{N}_{0_{\mathbb{Z}}}$ *defined by* $n \rightarrow \sim([s(n), 1])$ *forms a*
   a. *group isomorphism between* $\langle \mathbb{N}_0, + \rangle \rightarrow \langle \mathbb{N}_{0_{\mathbb{Z}}}, + \rangle$
   b. *group isomorphism between* $\langle \mathbb{N}_0, \cdot \rangle \rightarrow \langle \mathbb{N}_{0_{\mathbb{Z}}}, \cdot \rangle$
4. *For every* $z \in \mathbb{Z}$ *we have* $\exists x, y \in \mathbb{N}_{0_z}$ *such that* $z = x + (-y)$

As $0, 1 \in \mathbb{N}_{0_{\mathbb{Z}}}$ we have by 3.9 that $\langle \mathbb{N}_{0_{\mathbb{Z}}}, + \rangle$ and $\langle \mathbb{N}_{0_{\mathbb{Z}}}, \cdot \rangle$ are semi-groups

Proof.

1. We have the following
   a. $n, m \in \mathbb{N}_{0_{\mathbb{Z}}}$ then $\exists n', m' \in \mathbb{N}_0$ such that $n = \sim[(s(n'), 1)]$, $m = \sim[(s(m'), 1)]$ then $n+m = \sim[(s(n')+s(m'), 1+1)] = \sim[((n'+m'+1)+1, 1+1)] \underset{6.4}{=} \sim[(n'+m'+1, 1)] = \sim[(s(n'+m'), 1)] \in \mathbb{N}_{0_Z}$
   b. $0 = \sim[(1,1)] = \sim[(s(0), 1)] \in \mathbb{N}_{0_Z}$

   using 3.8 we have our proof.
2. We have the following
   a. $n, m \in \mathbb{N}_{0_{\mathbb{Z}}}$ then $\exists n', m' \in \mathbb{N}_0$ such that $n = \sim[(s(n'), 1)], m = \sim[(s(m'), 1)]$. we have then that $n \cdot m = \sim[(s(n') \cdot s(m') + 1 \cdot 1, 1 \cdot s(m') + 1 \cdot s(m'))] = \sim[s(n') \cdot s(m') + 1, s(m') + s(n')] \sim [(n'+1)(m'+1)+1, n'+1+m'+1] = \sim[(n' \cdot m' + n' + m' + 1, n' + m' + 1)] \underset{6.4}{=} \sim[(n' \cdot m' + 1, 1)] \in \mathbb{N}_{0_{\mathbb{Z}}}$
   b. $1 = \sim[(2,1)] = \sim[(s(1), 1)] \in \mathbb{N}_{0_{\mathbb{Z}}}$

   so using 3.8 we have our proof.
3. First we prove that the function $i_{\mathbb{N}_0}: \mathbb{N}_0 \rightarrow \mathbb{N}_{0_{\mathbb{Z}}}$ is a bijection
   a. **(Injectivity)** if $i_{\mathbb{N}_0}(n) = i_{\mathbb{N}_0}(m) \Rightarrow \sim[(s(n), 1)] = \sim[s(m), 1] \Rightarrow (s(n), 1) \sim (s(m), 1) \Rightarrow s(m) + 1 = s(n) + 1 \underset{4.44}{\Rightarrow} s(m) = s(n) \underset{4.15}{\Rightarrow} m = n$
   b. **(Surjectivity)** If $n \in \mathbb{N}_{0_{\mathbb{Z}}}$ then there exists a $n' \in \mathbb{N}_0 \vdash n = \sim[(s(n'), 1)] = i_{\mathbb{N}_0}(n')$

   Next we prove that $i_{\mathbb{N}_0}(n+m) = i_{\mathbb{N}_0}(n) + i_{\mathbb{N}_0}(m)$ and $i_{\mathbb{N}_0}(n \cdot m) = i_{\mathbb{N}_0}(n) \cdot i_{\mathbb{N}_0}(m)$
   a. $i_{\mathbb{N}_0}(n+m) = \sim[(s(n+m), 1)] \underset{6.4}{=} \sim[(s(n+m)+1, 1+1)] = \sim[(((n+m)+1)+1, 1+1)] = \sim[((n+1)+(m+1), 1+1)] = \sim[(n+1, 1)] + \sim[(m+1, 1)] = \sim[(s(n), 10)] + \sim[(s(m), 1)] = i_{\mathbb{N}_0}(n) + i_{\mathbb{N}_0}(m)$

b. $i_{\mathbb{N}_0}(n\cdot m)=\sim[(s(n\cdot m),1)]=\sim[(n\cdot m+1,1)]\underset{6.4}{=}\sim[(n\cdot m,0)]=\sim[(n,0)]\cdot\sim[(m,0)]\underset{6.4}{=}\sim[(n+1,1)]+\sim[(m+1,1)]=\sim[(s(n),1)]+\sim[(s(m),1)]=i_{\mathbb{N}_0}(n)\cdot i_{\mathbb{N}_0}(m)$

4. If $z\in\mathbb{Z}$ then $z=\sim[(n,m)]=\sim[(n,0)]+\sim[(0,m)]=\sim[(n,0)]+(-(\sim[(m,0)]))\underset{6.4}{=}\sim[(s(n),1)]+(-(\sim[(s(m),1)]))=x+(-y)$ where $x=\sim[(s(n),1)]\in\mathbb{N}_{0_\mathbb{Z}}$ and $y=\sim[(s(m),1)]\in\mathbb{N}_{0_\mathbb{Z}}$ □

Definition 6.23. $-\mathbb{N}_{0_\mathbb{Z}}=\{-n|n\in\mathbb{N}_{0_\mathbb{Z}}\}$, $-\mathbb{N}_{0_\mathbb{Z}}$ *is the collection of negative numbers.*

Lemma 6.24. $\mathbb{Z}=\mathbb{N}_{0_\mathbb{Z}}\bigcup(-\mathbb{N}_{0_\mathbb{Z}})$*, so the whole numbers is the union of the positive and negative numbers.*

Proof. First as $\mathbb{N}_{0_\mathbb{Z}}\subseteq\mathbb{Z}$ and thus $-\mathbb{N}_{0_\mathbb{Z}}\subseteq\mathbb{Z}$ we have $\mathbb{N}_{0_\mathbb{Z}}\bigcup(-\mathbb{N}_{0_\mathbb{Z}})\subseteq\mathbb{Z}$. Next assume that $z\in\mathbb{Z}$ then by 6.22 we have there exists a $x,y\in\mathbb{N}_{0_\mathbb{Z}}$ such that $z=x+(-y)$. Using 6.22 again there exists $x',y'\in\mathbb{N}_0$ with $x=i_{\mathbb{N}_0}(x')$ and $y=i_{\mathbb{N}_0}(y')$ where $i_{\mathbb{N}_0}$ is the isomorphism defined by $i_{\mathbb{N}_0}:\mathbb{N}_0\rightarrow\mathbb{N}_{0_\mathbb{Z}}$ $n\rightarrow f(n)=\sim[(s(n),1)]$. We have then the following cases to consider ($\mathbb{N}_0,\leqslant$ is fully ordered)

1. $x'=y'$ then $x=i_{\mathbb{N}_0}(x')=i_{\mathbb{N}_0}(y')=y\Rightarrow z=x+(-x)=0=\sim[(1,1)]=\sim[(s(0),1)]\in\mathbb{N}_{0_\mathbb{Z}}\Rightarrow z\in\mathbb{N}_{0_\mathbb{Z}}\bigcup(-\mathbb{N}_{0_\mathbb{Z}})$
2. $x'<y'$ then by 4.61 there exists a $k'\in\mathbb{N}$ such that $y'=x'+k'$. Take then $k=i_{\mathbb{N}_0}(k')\in\mathbb{N}_{0_\mathbb{Z}}$, by the linearity of $i_{\mathbb{N}_0}$ we have then $y=i_{\mathbb{N}_0}(y')=i_{\mathbb{N}_0}(x'+k')=i_{\mathbb{N}_0}(x')+i_{\mathbb{N}_0}(k')=x+k$ and thus $z=x+(-y)=x+(-1)\cdot y=x+(-1)\cdot(x+k)=x+((-x)+(-k))=-k\in-\mathbb{N}_{0_\mathbb{Z}}\Rightarrow z\in-\mathbb{N}_{0_\mathbb{Z}}\Rightarrow z\in\mathbb{N}_{0_\mathbb{Z}}\bigcup(-\mathbb{N}_{0_\mathbb{Z}})$
3. $y'<x'$ then by 4.61 there exists a $k'\in\mathbb{N}$ such that $x'=y'+k'$, if $k=i_{\mathbb{N}_0}(k')\in\mathbb{N}_{0_\mathbb{Z}}$ we have then $x=i_{\mathbb{N}_0}(x')=i_{\mathbb{N}_0}(y')+i_{\mathbb{N}_0}(k')=y+k$ and thus $z=x+(-y)=y+k+(-y)=k\in\mathbb{N}_{0_\mathbb{Z}}\Rightarrow z\in\mathbb{N}_{0_\mathbb{Z}}\Rightarrow z\in\mathbb{N}_{0_\mathbb{Z}}\bigcup(-\mathbb{N}_{0_\mathbb{Z}})$ □

We can now define a order relation on $\mathbb{Z}$.

Definition 6.25. (Order relation in $\mathbb{Z}$) *We define* $\leqslant\subseteq\mathbb{Z}\times\mathbb{Z}$ *by* $\leqslant=\{(n,m)|m+(-n)\in\mathbb{N}_{0_\mathbb{Z}}\}$ *in other words* $n\leqslant m$ *iff* $m+(-n)\in\mathbb{N}_{0_\mathbb{Z}}$

Theorem 6.26. $\langle\mathbb{Z},\leqslant\rangle$ *is a partially ordered set that is fully-ordered*

Proof.

1. **(reflectivity)** If $n\in\mathbb{Z}$ then $n+(-n)=0=\sim[(1,1)]=\sim[(s(0),1)]\in\mathbb{N}_{0_\mathbb{Z}}$
2. **(anti–symmetry)** If $n,m\in\mathbb{Z}$, $n\leqslant m$ and $m\leqslant n$ then $m+(-n)\in\mathbb{N}_{0_\mathbb{Z}}$ and $n+(-m)\in\mathbb{N}_{0_\mathbb{Z}}$. As $n,m\in\mathbb{Z}$ we have $n=\sim[(n_1,n_2)]$ and $m=\sim[(m_1,m_2)]$ then

$$\begin{aligned} n+(-m) &= \sim[(n_1+m_2,n_2+m_1)]=\sim[(s(k),1)]\text{ where }k\in\mathbb{N}_0\\ m+(-n) &= \sim[(m_1+n_2,m_2+n_1)]=\sim[(s(l),1)]\text{ where }l\in\mathbb{N}_0\end{aligned}$$

we have then

$$\begin{array}{rcl} n_1+m_2+1 & = & n_2+m_1+s(k) \\ & = & n_2+m_1+k+1 \\ & \Leftrightarrow & \\ n_1+m_2 & = & n_2+m_1+k \\ m_1+n_2+1 & = & m_2+n_1+s(l) \\ & = & m_2+n_1+l+1 \\ & \Leftrightarrow & \\ m_1+n_2 & = & m_2+n_1+l \end{array}$$

summing the two equalities we get

$$\begin{array}{rcl} n_1+m_2+m_1+n_2 & = & n_1+m_2+m_1+n_2+k+l \\ & \underset{4.44}{\Leftrightarrow} & \\ 0 & = & k+l \end{array}$$

Using 4.60 we have then that $k=l=0$ but then we have from $n_1+m_2=n_2+m_1+k$ that $n_1+m_2=n_2+m_1$ or $(n_1,n_2)\sim(m_1,m_2)\Rightarrow\sim[(n_1,n_2)]=\sim[(m_1,m_2)]\Rightarrow n=m$

3. **(transitivity)** If $n,m,k\in\mathbb{Z}$ then if $n\leqslant m$ and $m\leqslant k$ we have $m+(-n)\in\mathbb{N}_{0_\mathbb{Z}}$ and $k+(-m)\in\mathbb{N}_{0_\mathbb{Z}}$ then as we have that $\mathbb{N}_{0_\mathbb{Z}}$ is a sub-semi-group we have that $\mathbb{N}_{0_\mathbb{Z}}\ni(k+(-m))+(m+(-n))=k+((-m)+m)+(-n)=k+0+(-n)=k+(-n)\Rightarrow k\leqslant n$

4. **(fully ordering)** If $n,m\in\mathbb{Z}$ then by 6.24 we have for $n+(-m)\in\mathbb{Z}$ that either

    a. $n+(-m)\in\mathbb{N}_{0_\mathbb{Z}}\Rightarrow m\leqslant n$

    b. $n+(-m)\in-\mathbb{N}_{0_\mathbb{Z}}\Rightarrow-(n+(-m))\in\mathbb{N}_{0_\mathbb{Z}}\Rightarrow m+(-n)\in\mathbb{N}_{0_\mathbb{Z}}\Rightarrow n\leqslant m$ proving fully ordering. □

Theorem 6.27. *If $n,m\in\mathbb{Z}$ with $n\leqslant m$ (or $n<m$) then $-m\leqslant-n$ (or $-m<-n$*

Proof.

1. *If $n\leqslant m$ then $m+(-n)\in\mathbb{N}_{0_\mathbb{Z}}\Rightarrow(-n)+(-(-m))=(-n)+m=m+(-n)\in\mathbb{N}_{0_\mathbb{Z}}\Rightarrow-m\leqslant-n$*
2. *If $n<m\Rightarrow n\leqslant m\wedge n\neq m\Rightarrow-m\leqslant-n\wedge-n\neq-m\Rightarrow-m<-n$* □

Theorem 6.28. *If $n\in\mathbb{Z}$ then we have $n\in\mathbb{N}_{0_\mathbb{Z}}\Leftrightarrow 0\leqslant n$. Note that this is the motivation to call $\mathbb{N}_{0_\mathbb{Z}}$ the set of positive numbers.*

Proof.

$$\begin{array}{rcl} 0\leqslant z & \Leftrightarrow & n+(-0)\in\mathbb{N}_{0_\mathbb{Z}} \\ & \underset{6.13}{\Leftrightarrow} & n+(-1)\cdot 0\in\mathbb{N}_{0_\mathbb{Z}} \\ & \underset{4.37}{\Leftrightarrow} & n+0\in\mathbb{N}_{0_\mathbb{Z}} \\ & \Leftrightarrow & n\in\mathbb{N}_{0_\mathbb{Z}} \end{array}$$

□

Theorem 6.29. *If $n \in \mathbb{Z}$ so $n = \sim[(n_1, n_2)]$ then $0 < n$ iff $n_2 < n_1$*

Proof.

$\Rightarrow$

Assume that $0 < n$ then $n \neq 0$ and $0 \leqslant n$ so $0 \neq 0$ and $n \in \mathbb{N}_{0_{\mathbb{Z}}} \Rightarrow \exists m \in \mathbb{N}_0 \vdash n = \sim[(n_1, n_2)] \sim [(s(m), 1)] \Rightarrow n_1 + 1 = n_2 + s(m) = n_2 + (m+1) = (n_2 + m) + 1 \Rightarrow n_1 = n_2 + m$. Now if $m = 0 \Rightarrow n_1 = n_2 \Rightarrow n = \sim[(n_1, n_1)] \underset{6.7}{=} 0 \neq n$ a contradiction so we have $m \in \mathbb{N}$ but then by 4.61 we have $n_2 < n_1$.

$\Leftarrow$ If $n_2 < n_1 \underset{4.61}{\Rightarrow} \exists m \in \mathbb{N} \vdash n_1 = n_2 + m \Rightarrow n_1 + 1 = n_2 + s(m) \Rightarrow \sim[(n_1, n_2] = \sim[(s(m), 1)] \Rightarrow n \in \mathbb{N}_{0_{\mathbb{Z}}} \Rightarrow 0 \leqslant n$, now if $n = 0 = \sim[(1,1)]$ then $s(m) + 1 = 1 + 1 \Rightarrow m + (1+1) = 0 + (1+1) \Rightarrow m = 0$ a contradiction so $n \neq 0$ and thus $0 < n$ □

Corollary 6.30. *If $n \in \mathbb{Z}$ so $n = \sim[(n_1, n_2)]$ then $0 \leqslant n$ iff $n_2 \leqslant n_1$*

Proof.

$\Rightarrow$

If $0 \leqslant n$ then we have either

1. $n = 0 \underset{6.7}{\Rightarrow} n_1 = n_2 \Rightarrow n_2 \leqslant n_1$
2. $0 < n \underset{\text{previous theorem}}{\Rightarrow} n_2 < n_1 \Rightarrow n_2 \leqslant n_1$

$\Leftarrow$

If $n_2 \leqslant n_1$ then we have either

1. $n_2 = n_1 \Rightarrow n = \sim[(n_1, n_1)] = 0 \Rightarrow 0 \leqslant n$
2. $n_2 < n_1 \Rightarrow 0 < n \Rightarrow 0 \leqslant n$

□

Theorem 6.31. *$i_{\mathbb{N}_0}: \mathbb{N}_0 \rightarrow \mathbb{N}_{0_{\mathbb{Z}}}$ defined by $n \rightarrow i_{\mathbb{N}_0}(n) = \sim[(s(n), 1)]$ is order preserving, so if $n, m \in \mathbb{N}_0$ with $n \leqslant m$ then $i_{\mathbb{N}_0}(n) \leqslant i_{\mathbb{N}_0}(m)$*

Proof. If $n \leqslant m$ then we have the following cases

1. $n = m \Rightarrow i_{\mathbb{N}_0}(n) = i_{\mathbb{N}_0}(m) \Rightarrow \mathbb{N}_{0_{\mathbb{Z}}} \ni 0 = i_{\mathbb{N}_0}(m) + (-i_{\mathbb{N}_0}(n)) \Rightarrow i_{\mathbb{N}_0}(n) \leqslant i_{\mathbb{N}_0}(m)$
2. $n < m$ then $i_{\mathbb{N}_0}(m) + (-i_{\mathbb{N}_0}(n)) = \sim[(s(m), 1)] + (-(\sim[s(n), 1])) = \sim[(s(m), 1)] + \sim[(1, s(n))] = \sim[(s(m)+1, s(n)+1)] = \sim[(s(m), s(n))] = \sim[(m+1, n+1)] = \sim[(m, n)]$. By 6.29 we have from $n < m$ that $0 < \sim[(m, n)] = i_{\mathbb{N}_0}(m) + (-i_{\mathbb{N}_0}(n)) \Rightarrow 0 \leqslant i_{\mathbb{N}_0}(m) + (-i_{\mathbb{N}_0}(n)) \underset{6.28}{\Rightarrow} i_{\mathbb{N}_0}(m) + (-i_{\mathbb{N}_0}(n)) \in \mathbb{N}_{0_{\mathbb{Z}}} \Rightarrow i_{\mathbb{N}_0}(n) \leqslant i_{\mathbb{N}_0}(m)$ □

Theorem 6.32. *If $n, m, k \in \mathbb{Z}$ then $n < m \Rightarrow n + k < m + k$*

Proof. From $n < m \Rightarrow n \neq m$ and $n \leqslant m$ we have $n + (-m) \in \mathbb{N}_{0_{\mathbb{Z}}} \Rightarrow n + 0 + (-m) \in \mathbb{N}_{0_{\mathbb{Z}}} \Rightarrow n + k + (-k) + (-m) \in \mathbb{N}_{0_{\mathbb{Z}}} \Rightarrow (n+k) + (-(m+k)) \in \mathbb{N}_{0_{\mathbb{Z}}} \Rightarrow n + k \leqslant m + k$. If now $n + k = m + k \Rightarrow n + k + (-k) = m + k + (-k) \Rightarrow n + 0 = m + 0 \Rightarrow n = m$ a contradiction with $n < m$ so we have $n + k < m + k$. □

Corollary 6.33. *If* $n, m, k \in \mathbb{Z}$ *then* $n \leqslant m \Rightarrow n + k \leqslant m + k$

Proof. If $n \leqslant m$ then we have either

1. $n = m \Rightarrow n + k = m + k \Rightarrow n + k \leqslant m + k$
2. $n < m \underset{\text{previous theorem}}{\Rightarrow} n + k < m + k \Rightarrow n + k \leqslant m + k$ □

Theorem 6.34. *If* $n, m \in \mathbb{Z}$ *then* $n < m \Leftrightarrow 0 < m + (-n)$

Proof.

$\Rightarrow$

If $n < m \Rightarrow n \leqslant m$ and $n \neq m$ so $n + (-m) \in \mathbb{N}_{0_{\mathbb{Z}}} \underset{6.28}{\Rightarrow} 0 \leqslant n + (-m)$. Now if $0 = n + (-m) \Rightarrow 0 + m = n + (-m) + m \Rightarrow m = n + 0 = n$ contradicting $n < m$.

$\Leftarrow$

If $0 < m + (-n)$ then by the previous corollary we have $0 + n < m + (-n) + n \Rightarrow n < m + 0 \Rightarrow n < m$

□

Corollary 6.35. *If* $n, m \in \mathbb{Z}$ *then* $n \leqslant m \Leftrightarrow 0 \leqslant m + (-n)$

Proof.

$$\begin{aligned} n \leqslant m &\Leftrightarrow n + (-m) \in \mathbb{N}_{0_{\mathbb{Z}}} \\ &\Leftrightarrow 0 \leqslant n + (-m) \end{aligned}$$

□

Theorem 6.36. *If* $n, m \in \mathbb{Z}$ *with* $0 < n$ *and* $0 < m \Rightarrow 0 < n \cdot m$

Proof. If $n = \sim[(n_1, n_2)], m = \sim[(m_1, m_2)]$ then from $0 < n$ and $0 < m$ we have by 6.29 $n_2 < n_1$ and $m_2 < m_1 \Rightarrow \exists k, l \in \mathbb{N} \vdash n_1 = n_2 + k$ and $m_1 = m_2 + l$ this gives

$$\begin{aligned} n_1 \cdot m_1 + n_2 \cdot m_2 &= (n_2 + k) \cdot (m_2 + l) + n_2 \cdot m_2 \\ &= n_2 \cdot m_2 + n_2 \cdot m_2 + n_2 \cdot l + m_2 \cdot k + k \cdot l \\ n_2 \cdot m_1 + n_1 \cdot m_2 &= n_2 \cdot (m_2 + l) + (n_2 + k) \cdot m_2 \\ &= n_2 \cdot m_2 + n_2 \cdot m_2 + n_2 \cdot l + m_2 \cdot k \end{aligned}$$

So using 4.61 and ($0 < k$ then by 4.70 and $0 < l$ we have $0 < k \cdot l$) we have that $n_2 \cdot m_1 + n_1 \cdot m_2 < n_1 \cdot m_1 + n_2 \cdot m_2$. Now $n \cdot m = \sim[(n_1 \cdot m_1 + n_2 \cdot m_2, n_2 \cdot m_1 + n_1 \cdot m_2)]$ and thus we have $0 < n \cdot m$ □

Theorem 6.37. *If* $n, m \in \mathbb{Z}$ *and* $k \in \mathbb{Z} \vdash k > 0$ *then* $n < m \Rightarrow n \cdot k < m \cdot k$

Proof. From $n < m$ we have by 4.56 that $0 < m + (-n)$ then by the above theorem we have $0 < (m + (-n)) \cdot k = m \cdot k + (-1) \cdot n \cdot k = m \cdot k + (-1)(n \cdot k) = m \cdot k + (-(n \cdot k)) \underset{6.34}{\Rightarrow} n \cdot k < m \cdot k$ □

Corollary 6.38. *If* $n, m \in \mathbb{Z}$ *and* $k \in \mathbb{Z} \vdash k \geqslant 0$ *then* $n \leqslant m \Rightarrow n \cdot k \leqslant m \cdot k$

Proof. For k we have the following possibilities

1. $k = 0 \Rightarrow n \cdot k = 0 = m \cdot k \Rightarrow n \cdot k \leqslant m \cdot k$

2. $0<k$ then for $n\leqslant m$ we have either

   a. $n=m\Rightarrow n\cdot k=m\cdot k\Rightarrow n\cdot k\leqslant m\cdot k$

   b. $n<m\underset{\text{previous theorem}}{\Rightarrow} n\cdot k<m\cdot k=n\cdot k\leqslant m\cdot k$ □

Theorem 6.39. *If $n\in\mathbb{Z}$ then $0\leqslant n\cdot n$*

Proof. Let $n=\sim[(n_1,n_2)]$ then $n\cdot n=\sim[(n_1\cdot n_1+n_2\cdot n_2,n_2\cdot n_1+n_1\cdot n_2)]$. From the totality of order in $\mathbb{N}_0$ we have either

1. $n_1=n_2$ then $n=\sim[(n_1,n_1)]=0\underset{6.12}{\Rightarrow} n\cdot n=0\cdot 0=0\Rightarrow 0\leqslant n\cdot n$
2. $n_1<n_2$ then we have $\exists k\in\mathbb{N}$ with $n_2=n_1+k$

$$\begin{array}{rcl} n_1\cdot n_1+n_2\cdot n_2 &=& n_1\cdot n_1+(n_1+k)\cdot(n_1+k)\\ &=& (n_1\cdot n_1+n_1\cdot n_1+n_1\cdot k+n_1\cdot k)+k\cdot k\\ n_2\cdot n_1+n_1\cdot n_2 &=& (n_1+k)\cdot n_1+n_1\cdot(n_1+k)\\ &=& (n_1 n_1+n_1\cdot n_1+n_1\cdot k+n_1\cdot k) \end{array}$$

   So as $0<k\cdot k$ we have $n_2\cdot n_1+n_1\cdot n_2<n_1\cdot n_1+n_2\cdot n_2\Rightarrow 0<n\cdot n\Rightarrow 0\leqslant n\cdot n$

3. $n_2<n_1$ then we have $\exists k\in\mathbb{N}$ with $n_1=n_2+k$

$$\begin{array}{rcl} n_1\cdot n_1+n_2\cdot n_2 &=& (n_2+k)\cdot(n_2+k)+n_2\cdot n_2\\ &=& (n_2\cdot n_2+n_2\cdot n_2+n_2\cdot k+n_2\cdot k)+k\cdot k\\ n_2\cdot n_1+n_1\cdot n_2 &=& n_2\cdot(n_2+k)+(n_2+k)\cdot n_2\\ &=& (n_2 n_2+n_2\cdot n_2+n_2\cdot k+n_2\cdot k) \end{array}$$

   So as $0<k\cdot k$ we have $n_2\cdot n_1+n_1\cdot n_2<n_1\cdot n_1+n_2\cdot n_2\Rightarrow 0<n\cdot n\Rightarrow 0\leqslant n\cdot n$

□

Theorem 6.40. (Archimedean property of $\mathbb{Z}$) *If $m,n\in\mathbb{Z}$ with $m>0$ then there exists a $k\in\mathbb{N}_{0_{\mathbb{Z}}}$ with $n<k\cdot m$*

Proof. We have the following cases for n

1. $n\leqslant 0$ then if we take $k=1\in\mathbb{N}_{0_{\mathbb{Z}}}$ then $n\leqslant 0<m=1\cdot m=k\cdot m\Rightarrow n<k\cdot m$
2. If $n>0\Rightarrow n\in\mathbb{N}_{0_{\mathbb{Z}}}$ as also $m>0$ we have $m\in\mathbb{N}_{0_{\mathbb{Z}}}$ and there exists a $n',m'\in\mathbb{N}_0$ such that $n=i_{\mathbb{Z}}(n')$, $m=i_{\mathbb{Z}}(m')$. Assume now that $m'\leqslant 0\Rightarrow i_{\mathbb{Z}}(m')\leqslant i_{\mathbb{Z}}(0)=\sim[(s(0),1)]=\sim[(1,1)]=0$ a contradiction so we must have that $m'>0$. Then by 4.73 there exists a $k'\in\mathbb{N}_0$ such that $n'<k'\cdot m'\Rightarrow n=i_{\mathbb{Z}}(n')<i_{\mathbb{Z}}(k'\cdot m')=i_{\mathbb{Z}}(k')\cdot i_{\mathbb{Z}}(m')=k\cdot m$ where $k=i_{\mathbb{Z}}(k')\in\mathbb{N}_{0_{\mathbb{Z}}}$ □

Theorem 6.41. *$\langle\mathbb{N}_{0_{\mathbb{Z}}},\leqslant\rangle$ is well-ordered.*

Proof. Let $A\subseteq\mathbb{N}_{0_{\mathbb{Z}}}$ be a non empty subset of $\mathbb{N}_{0_{\mathbb{A}}}$ then as $i_{\mathbb{N}_0}:\mathbb{N}_0\rightarrow\mathbb{N}_{0_{\mathbb{Z}}}$ is a bijection we have that $i_{\mathbb{N}_0}^{-1}(A)$ is non empty and thus by the well-ordering of the natural numbers (see 4.52) there exists a $m'=\min(i_{\mathbb{N}_0}^{-1}(A))\in i_{\mathbb{N}_0}^{-1}(A)$. Take now $m=i_{\mathbb{N}_0}(m')$ then we have

1. $m'\in i_{\mathbb{N}_0}^{-1}(A)\Rightarrow m=i_{\mathbb{N}_0}(m')\in A$

2. If $a \in A \underset{i_{\mathbb{N}_0}\text{is bijective}}{\Rightarrow} \exists a' = i_{\mathbb{N}_0}^{-1}(a) \in i_{\mathbb{N}_0}^{-1}(A) \underset{m'\text{is least element of }A}{\Rightarrow} m' \leqslant a' \underset{i_{\mathbb{N}_0}\text{is order preserving (see 6.31)}}{\Rightarrow} m = i_{\mathbb{N}_0}(m') \leqslant i_{\mathbb{N}_0}(a') = a \Rightarrow m = \min(A)$ □

Corollary 6.42. $\langle \mathbb{N}_{0_\mathbb{Z}}, \leqslant \rangle$ *is conditionally complete (any non empty subset in* $\mathbb{N}_{0_\mathbb{Z}}$ *bounded above has a supremum (lowest upper bound). Furthermore if* $\emptyset \neq A \subseteq \mathbb{N}_0$ *then if* $\sup(A)$ *exists we have* $\sup(A) \in A$

Proof. The fact that $\langle \mathbb{N}_{0_\mathbb{Z}}, \leqslant \rangle$ is conditionally complete follows from the fact that $\langle \mathbb{N}_{0_\mathbb{Z}}, \leqslant \rangle$ is well-ordered (see 6.41) and 2.184. Now if $\emptyset \neq A \subseteq \mathbb{N}_{0_\mathbb{Z}}$ then there exists a $a \in A \Rightarrow i_{\mathbb{N}_0}^{-1}(a) \in i_{\mathbb{N}_0}^{-1}(A) \Rightarrow \emptyset \neq i_{\mathbb{N}_0}^{-1}(A)$. Also if $M = \sup(A)$ take then $m = i_{\mathbb{N}_0}^{-1}(M)$. If now $a \in A \Rightarrow a \leqslant m$ [For if $m < a \underset{i_{\mathbb{N}_0}\text{is order preserving and injective}}{\Rightarrow} M = i_{\mathbb{N}_0}(m) < i_{\mathbb{N}_0}(a) \in A$ contradicting the fact that M is a upper bound of A]. Furthermore if m' is another upper bound of $i_{\mathbb{N}_0}^{-1}(A)$ then if $a \in A$ we have $i_{\mathbb{N}_0}^{-1}(a) \in i_{\mathbb{N}_0}^{-1}(A) \Rightarrow i_{\mathbb{N}_0}^{-1}(a) \leqslant m' \underset{i_{\mathbb{N}_0}^{-1}\text{is order preserving}}{\Rightarrow} a = i_{\mathbb{N}_0}(i_{\mathbb{N}_0}^{-1}(a)) \leqslant i_{\mathbb{N}_0}(m') \Rightarrow i_{\mathbb{N}_0}(m')$ is a upper bound of A and as M is the least upper bound we have $M \leqslant i_{\mathbb{N}_0}(m')$ and then $i_{\mathbb{N}_0}^{-1}(M) \leqslant m'$ [otherwise is $m' < i_{\mathbb{N}_0}^{-1}(M) \underset{i_{\mathbb{N}_0}\text{is order preserving}}{\Rightarrow} i_{\mathbb{N}_0}(m') < i_{\mathbb{N}_0}(i_{\mathbb{N}_0}^{-1}(M)) = M \leqslant i_{\mathbb{N}_0}(m') \Rightarrow i_{\mathbb{N}_0}(m') < i_{\mathbb{N}_0}(m')$ a contradiction] so $m = i_{\mathbb{N}_0}^{-1}(M)$ is the lowest upper bound of $i_{\mathbb{N}_0}^{-1}(A)$ or in other words $m = \sup(i_{\mathbb{N}_0}^{-1}(A))$, using 4.67 we have then that $i_{\mathbb{N}_0}^{-1}(M) = m \in i_{\mathbb{N}_0}^{-1}(A) \Rightarrow M = i_{\mathbb{N}_0}(i_{\mathbb{N}_0}^{-1}(M)) \in A \Rightarrow \sup(A) \in M$ □

Lemma 6.43. $\forall n, m \in \mathbb{N}_{0_Z}$ *then* $n < m \Leftrightarrow \exists k \in \mathbb{N}_{0_\mathbb{Z}} \backslash \{0\}$ *such that* $m = n + k$

Proof.

(⇒)

Take $n' = i_{\mathbb{N}_0}^{-1}(n)$ and $m' = i_{\mathbb{N}_0}^{-1}(m)$ then $n' < m'$ [otherwise if $m' \leqslant n'$ we have by the order preserving properties of $i_{\mathbb{N}_0}$ that $m = i_{\mathbb{N}_0}(i_{\mathbb{N}_0}^{-1}(m)) \leqslant i_{\mathbb{N}_0}(i_{\mathbb{N}_0}^{-1}(n)) = n \Rightarrow m \leqslant n < m \Rightarrow m < m$ a contradiction]. Using 4.61 there exists a $k' \in \mathbb{N}_0 \backslash \{0\}$ such that $m' = n' + k'$ then using the fact that $i_{\mathbb{N}_0}$ is a isomorphism we have $m = i_{\mathbb{N}_0}(i_{\mathbb{N}_0}^{-1}(m)) = i_{\mathbb{N}_0}(m') = i_{\mathbb{N}_0}(n' + k') = i_{\mathbb{N}_0}(n') + i_{\mathbb{N}_0}(k') = i_{\mathbb{N}_0}(i_{\mathbb{N}_0}^{-1}(n)) + k = n + k$ where $k = i_{\mathbb{N}_0}(k') \in \mathbb{N}_{0_\mathbb{Z}}$ and $k \neq 0$ [otherwise if $k = 0 \Rightarrow i_{\mathbb{N}_0}(0) = 0 = k = i_{\mathbb{N}_0}(k') \underset{i_{\mathbb{N}_0}\text{is a bijection}}{\Rightarrow} k' = 0$ contradicting $k' \in \mathbb{N}_0 \backslash \{0\}$]

(⇐)

If $m = n + k$ where $k \in \mathbb{N}_{0_Z} \backslash \{0\} \Rightarrow 0 < k = (n + k) + (-n) = m + (-n) \Rightarrow 0 < m + (-n) \Rightarrow n = 0 + n < m + (-n) + n = m \Rightarrow n < m$ □

Definition 6.44. *Given* $z \in \mathbb{Z}$ *then* $|z| \in \mathbb{N}_{0_\mathbb{Z}}$ *is defined by* $|z| = \begin{cases} z \text{ if } 0 \leqslant z \\ -z \text{ if } z < 0 \end{cases}$. *So if* $z \geqslant 0 \Rightarrow z = |z|$

Note 6.45. Given $z, z' \in \mathbb{Z}$ then its easy to prove that $|z \cdot z'| = |z| \cdot |z'|$

Proof. We have the following cases for z and z' to consider

1. $(\mathbf{0} \leqslant \boldsymbol{z} \wedge \mathbf{0} \leqslant \boldsymbol{z'})$ then $0 \leqslant z \cdot z'$ and $|z| \cdot |z'| = z \cdot z' = |z \cdot z'|$
2. $(\mathbf{0} \leqslant \boldsymbol{z} \wedge \boldsymbol{z'} < \mathbf{0})$ then $z \cdot z' \leqslant 0$ and $z \cdot z' = -|z \cdot z'| \Rightarrow -(z \cdot z') = |z \cdot z'|$, also $-(z \cdot z') = z \cdot (-z') = |z| \cdot |z'| \Rightarrow |z| \cdot |z'| = |z \cdot z'|$

3. $\boldsymbol{(z<0 \wedge 0 \leqslant z')}$ then $z \cdot z' \leqslant 0$ and $z \cdot z' = -|z \cdot z'| \Rightarrow -(z \cdot z') = |z \cdot z'|$, also $-(z \cdot z') = (-z) \cdot z' = |z| \cdot |z'| \Rightarrow |z| \cdot |z'| = |z \cdot z'|$
4. $\boldsymbol{(z<0 \wedge z'<0)}$ then $0 \leqslant -z \wedge 0 \leqslant -z' \Rightarrow 0 \leqslant (-z) \cdot (-z') = z \cdot z' \Rightarrow |z \cdot z'| = z \cdot z' = (-|z|) \cdot (-|z'|) = |z| \cdot |z'|$ □

Lemma 6.46. *If $m \in \mathbb{Z}$ is such that $0<m \Rightarrow 1 \leqslant m$*

Proof. If $m \in \mathbb{Z}$ is such that $0<m$ then using 6.28 we have $m \in \mathbb{N}_{0_{\mathbb{Z}}}$. We can not have $i_{\mathbb{N}_0}^{-1}(m)=0$ [for then $m = i_{\mathbb{N}_0}(i_{\mathbb{N}_0}^{-1}(m)) = i_{\mathbb{N}_0}(0) = \sim[(s(0),1)] = \sim[(1,1)] = 0]$ and thus $i_{\mathbb{N}_0}^{-1}(m) \neq 0$. Using 4.47 we have then $0 < i_{\mathbb{N}_0}^{-1}(m)$ which by 4.51 means that $1 = s(0) \leqslant i_{\mathbb{N}_0}^{-1}(m) \Rightarrow 1 \leqslant i_{\mathbb{N}_0}^{-1}(m)$. From the order preserving of $i_{\mathbb{N}_0}$ we have then $i_{\mathbb{N}_0}(1) \leqslant i_{\mathbb{N}_0}(i_{\mathbb{N}_0}^{-1}(m)) = m \Rightarrow \sim[(s(1),1)] \leqslant m \Rightarrow 1 = \sim[(2,1)] \leqslant m \Rightarrow 1 \leqslant m$ □

Theorem 6.47. *If $n, m \in \mathbb{Z}$, $m \neq 0$ and $n|m$ then $n \leqslant |m|$*

Proof. From $n|m$ there exists a q such that $n \cdot q = m$ and as $m \neq 0$ we must have $q \neq 0$ (otherwise $m = n \cdot 0 = 0$). Then we have the following cases to consider

1. $\boldsymbol{(0<m \wedge n \leqslant 0)}$ in this case we have trivially $n < m = |m| \Rightarrow n \leqslant |m|$
2. $\boldsymbol{(0<m \wedge 0<n)}$ we must have $0<q$ [if $q \leqslant 0 \underset{q \neq 0}{\Rightarrow} q < 0 \underset{0<n}{\Rightarrow} m = n \cdot q < n \cdot 0 = 0 \Rightarrow m < 0 < m \Rightarrow m < m$ a contradiction] so using the previous lemma we have then $1 \leqslant q \underset{0<n}{\Rightarrow} n = n \cdot 1 \leqslant n \cdot q = m = |m| \Rightarrow n \leqslant |m|$
3. $\boldsymbol{(m<0 \wedge n \leqslant 0)}$ here $0 < -m = |m| \Rightarrow n < |m| \Rightarrow n \leqslant |m|$
4. $\boldsymbol{(m<0 \wedge 0<n)}$ then $q<0$ [otherwise if $0 \leqslant q \underset{q \neq 0}{\Rightarrow} 0 < q \underset{0<n}{\Rightarrow} 0 < n \cdot q = m \Rightarrow 0 < 0$ a contradiction] and we have $0 < -q \underset{\text{previous lemma}}{\Rightarrow} 1 \leqslant -q \underset{0<n}{\Rightarrow} n = n \cdot 1 \leqslant n \cdot (-q) = -(n \cdot q) = -m = |m| \Rightarrow n \leqslant |m|$ □

Theorem 6.48. (Division Algorithm) *If $m, n \in \mathbb{Z}$ and $n>0$ then there exists a unique $r \in \mathbb{N}_0$ with $0 \leqslant r < n$ and a unique $q \in \mathbb{Z}$ such that $m = n \cdot q + r$*

Proof. Let $A_{m,n} = \{m + n \cdot q | q \in \mathbb{Z} \wedge 0 \leqslant m + n \cdot q\} \subseteq \mathbb{N}_{0_{\mathbb{Z}}}$. Using the Archimedean property of $\mathbb{Z}$ (see 6.40) there exists a $k \in \mathbb{N}_{0_{\mathbb{Z}}}$ such that $k \cdot n > -m \Rightarrow m + n \cdot k > 0 \Rightarrow m + n \cdot k \in A_{m,n} \Rightarrow A_{m,n} \neq \emptyset$. By the fact that $\langle \mathbb{N}_{0_{\mathbb{Z}}}, \leqslant \rangle$ is well-ordered we have then that $r' = \min(A_{m,n})$ exists. From $r' \in A_{m,n}$ we have then also $0 \leqslant r'$ and a $q' \in \mathbb{N}_0$ such that $r' = m + n \cdot q'$. For the relation between r' and n we have then the following possibilities

1. $\boldsymbol{(r'=n)}$ in this case we have $m + n \cdot q' = r' = n \Rightarrow m = n \cdot (1 - q')$, thus taking $q = (1 - q')$ and $r = 0 \Rightarrow 0 \leqslant r < n$ we have $m = n \cdot q + r$
2. $\boldsymbol{(r'<n)}$ then $r' = m + n \cdot q' \Rightarrow m = n \cdot (-q') + r'$ so if $r = r' \Rightarrow 0 \leqslant r < n$ and $q = -q'$ then $m = n \cdot q + r$
3. $\boldsymbol{(n<r')}$ using 6.43 there exists a $k \in \mathbb{N}_{0_{\mathbb{Z}}} \setminus \{0\} \Rightarrow 0 < k$ such that $r' = n + k \Rightarrow n + k = m + n \cdot q' \Rightarrow 0 < k = m + n \cdot (q' - 1) \Rightarrow k \in A_{\mathrm{mn}}$. From $r' = n + k \underset{0<n \text{ and } 6.43}{\Rightarrow} k < r' \Rightarrow r' \neq \min(A_{m,n})$ a contradiction. So this case does not apply.

Now to prove uniqueness assume that there exists another $q'', r'' \in \mathbb{Z}$ with $0 \leqslant r'' < n$ and $m = n \cdot q'' + r''$. We have then that $n \cdot q'' + r'' = n \cdot q + r \Rightarrow n \cdot (q - q'') = r'' - r$. We prove now by contradiction that $r = r''$, so assume that $r \neq r''$ then we have either

1. $(\boldsymbol{r < r''})$ then $0 < r'' - r = n \cdot (q - q'')$ which as $n > 0$ means that $0 < q - q''$ [otherwise from $q - q' \leqslant 0$ we would have $n \cdot (q - q'') \leqslant 0$ contradicting $0 < n \cdot (q - q'')$] so using 6.46 we have $1 \leqslant q - q''$. Also from $0 \leqslant r, r'' < n$ we have $r'' - r < n - r \leqslant n \Rightarrow n \cdot (q - q'') < n \Rightarrow q - q'' < 1$ [if $1 \leqslant q - q'' \underset{n>0}{\Rightarrow} n = n \cdot 1 \leqslant n \cdot (q - q'') < n \Rightarrow n < n$ a contradiction]. So we finally reach the conclusion that $1 \leqslant q - q'' < 1 \Rightarrow 1 < 1$ a contradiction.
2. $(\boldsymbol{r'' < r})$ then $0 < r - r'' = n \cdot (q'' - q)$ which as $n > 0$ means that $0 < q'' - q$ [otherwise from $q'' - q \leqslant 0$ we would have $n \cdot (q'' - q) \leqslant 0$ contradicting $0 < n \cdot (q'' - q)$] so using 6.46 we have $1 \leqslant q'' - q$. Also from $0 \leqslant r, r'' < n$ we have $r - r'' < n - r'' \leqslant n \Rightarrow n \cdot (q'' - q) < n \Rightarrow q'' - q < 1$ [if $1 \leqslant q'' - q \underset{n>0}{\Rightarrow} n = n \cdot 1 \leqslant n \cdot (q'' - q) < n \Rightarrow n < n$ a contradiction]. So we finally reach the conclusion that $1 \leqslant q'' - q < 1 \Rightarrow 1 < 1$ a contradiction. □

Definition 6.49. *Given $n, m \in \mathbb{Z}$ then we say that n divides m, noted by $n|m$ if there exists a $q \in \mathbb{Z}$ such that $q \cdot n = m$*

Note 6.50. If $n|m$ then we have also $(-n)|m$ and this proves that then also $(|d|)|m$

Proof. If $n|m$ then there exists a $q \in \mathbb{Z}$ such that $m = n \cdot q = (-n) \cdot (-q)$ where $-q \in \mathbb{Z}$ so that $(-n)|m$ □

Definition 6.51. *If $n, m \in \mathbb{Z}$ and $n|m$ then there exists a unique (see 6.48) $q \in \mathbb{Z}$ such that $n \cdot q = m$. This number is called the quotient of m and n and noted by m/n. So if $n|m$ then there*

Definition 6.52. *Given $n, m \in \mathbb{Z}$ then d is a common divisor for n and m if $d|n$ and $d|m$. The greatest common divisor of m and n noted by $\gcd(n, m) = \sup(D_{n,m}) \in \mathbb{N}_{0_\mathbb{Z}} \backslash \{0\}$ if it exists, where $D_{n,m} = \{d \in \mathbb{N}_{0_\mathbb{Z}} \backslash \{0\} | d|n \wedge d|m\}$. Note that $1 \in D_{n,m} \Rightarrow D_{n,m} \neq \emptyset$, note also that from 6.42 we have that if $\gcd(n, m) = \sup(D_{n,m})$ exists then $\gcd(n, m) \in D_{n,m} \Rightarrow \gcd(n, m)|n$ and $\gcd(n, m)|m$*

Theorem 6.53. *Given $n, m \in \mathbb{Z}$ then $\gcd(n, m)$ exists $\Leftrightarrow n \neq 0$ or $m \neq 0$*

Proof. Define $D_{n,m} = \{d \in \mathbb{Z} | d|n \wedge d|m\}$ so that $\gcd(n, m) = \sup(D_{n,m})$. The proof is delivered in two parts

1. $(\Rightarrow)$ So assume that $\gcd(n, m)$ exists and assume that $n = 0 = m$. Then if $q \in \mathbb{Z}$ we have $q \cdot 0 = 0 = n = m \Rightarrow q|n \wedge q|m \Rightarrow \mathbb{Z} = D_{n,m}$. Now from $0 < 1 \underset{\text{add} \gcd(n,m)}{\Rightarrow} \gcd(n, m) < \gcd(n, m) + 1 \in D_{n,m}$ contradicting the fact that $\gcd(n, m)$ is a upper bound of $D_{n,m}$.
2. $(\Leftarrow)$ Suppose that $n \neq 0$ or $m \neq 0$ then we have the following cases
   a. $(\boldsymbol{n \neq 0})$ from 6.47 we have then if $d \in D_{n,m}$ then $d|n \Rightarrow d < |n|$ so $D_{n,m}$ has a upper bound and by conditional completeness (see 6.42) we have that $\sup(D_{n,m})$ exists.

b. **($m \neq 0$)** from 6.47 we have then if $d \in D_{n,m}$ then $d|m \Rightarrow d < |m|$ so $D_{n,m}$ has a upper bound and by conditional completeness (see 6.42) we have that $\sup(D_{n,m})$ exists. □

Theorem 6.54. *If $n, m \in \mathbb{Z}$ with $m \neq 0$ then by the above theorem $\gcd(n,m)$ exists and from the fact that $\gcd(n,m)|m$, $\gcd(n,m)|n$ we can consider $n/\gcd(n,m)$ and $m/\gcd(n,m)$. We have now that if $d|(n/\gcd(n,m))$ and $d|(m/\gcd(n,m))$ then $d = 1$ or $d = -1$ furthermore $\gcd(n/\gcd(n,m), m/\gcd(n,m)) = 1$*

Proof. Take $n' = n/\gcd(n,m)$ and $m' = n/\gcd(n,m)$ then we have $n' \cdot \gcd(n, m) = n$ and $m' \cdot \gcd(n,m) = m$. If $d|n'$ then there exists a n'' such that $n'' \cdot d = n' \Rightarrow n'' \cdot d \cdot \gcd(n,m) = n' \cdot \gcd(n,m) = n$ or $(d \cdot \gcd(n,m))|n \underset{0<\gcd(n,m)}{\Rightarrow} (|d| \cdot \gcd(n, m))|n|$. Also if $d|m'$ then there exists a m'' such that $m'' \cdot d = m' \Rightarrow m'' \cdot d \cdot \gcd(n, m) = m' \cdot \gcd(n,m) = m$ or $(d \cdot \gcd(n,m))|m \underset{0<\gcd(n,m)}{\Rightarrow} (|d| \cdot \gcd(n,m))|m$. Now $0 \leqslant |d|$ and as $d \neq 0$ [otherwise $m = m'' \cdot d \cdot \gcd(n,m) = 0$ contradicting $m \neq 0$] we have $0 < |d|$ and as $0 < \gcd(n,m)$ we have $0 < |d| \cdot \gcd(n,m)$ and thus $|d| \cdot \gcd(n, m) \in D_{n,m} \Rightarrow |d| \cdot \gcd(n,m) \leqslant \gcd(n,m)$ and thus $|d| \leqslant 1$ [otherwise $\gcd(n,m) < |d| \cdot \gcd(n,m)$]. As we have also by 6.46 and $0 < |d|$ that $1 \leqslant |d|$ we must conclude that $|d| = 1$ and thus $d = 1$ or $d = -1$. As $D_{n',m'} = \{d \in \mathbb{N}_{0_{\mathbb{Z}}} \setminus \{0\} | d|n' \wedge d|m'\} = \{1\}$ we have $\gcd(n', m') = 1$. □

Definition 6.55. *Given $z \in \mathbb{Z}$ then we say z is even if $2|z$. If z is not even then we call Z odd.*

Theorem 6.56. *Given $z \in \mathbb{Z}$ then we have*

1. *z is even $\Leftrightarrow \exists m \in \mathbb{Z} \vdash z = 2 \cdot m$*
2. *z is odd $\Leftrightarrow \exists m \in \mathbb{Z} \vdash z = 2 \cdot m + 1$*
3. *z can not be even and odd*

Proof.

1. This follows directly from the definition of $2|z$
2. Using the division algorithm (see 6.48) there exists unique $n, m \in \mathbb{Z}$ with $z = 2 \cdot n + m$ where $0 \leqslant m < 2$. If now $1 < m \Rightarrow 0 < m - 1 \underset{6.46}{\Rightarrow} 1 \leqslant m - 1 \Rightarrow 2 \leqslant m < 2 \Rightarrow 2 < 2$ a contradiction so we have $0 \leqslant m \leqslant 1$. We have then

   a. **(z is odd)** then z is not even so we must have $0 < m \Rightarrow 1 \leqslant m \leqslant 1 \Rightarrow m = 1 \Rightarrow z = 2 \cdot m + 1$

   b. **($z = 2 \cdot m + 1$)** from the uniqueness of n, m we can not have $z = 2 \cdot m' + 0$ because we would have then the contradiction $1 = 0$

3. If z is even and odd then $\exists m, m' \in \mathbb{Z}$ such that $2 \cdot m = z = 2 \cdot m' + 1$ hence $2 \cdot (m - m') = 1$ we have now the following 3 cases to consider

   $m = m'$. then $2 \cdot 0 = 1 \Rightarrow 0 = 1$ a contradiction

   $m < m'$. then $m + 1 \leqslant m' \Rightarrow m - m' \leqslant -1 \Rightarrow 2 \cdot (m - m') \leqslant -2 \Rightarrow 1 \leqslant -2$ a contradiction

$\boldsymbol{m'<m}$. then $m'+1\leqslant m\Rightarrow 1\leqslant m-m'\Rightarrow 2\leqslant 2\cdot(m-m')=1\Rightarrow 2\leqslant 1$ a contradiction

so in all cases we have a contradiction. □

Theorem 6.57. *Given $m\in\mathbb{Z}$ such that $z\cdot z$ is even then we have that z is even*

Proof. We prove this by contradiction, so assume that z is odd then by the above theorem we have $z=2\cdot n+1$ and thus $z\cdot z=(2\cdot n+1)\cdot(2\cdot n+1)=2\cdot 2\cdot n\cdot n+2\cdot n\cdot 1+1\cdot 2\cdot n+1\cdot 1=2\cdot(2\cdot n)+2\cdot n+2\cdot n+1=2\cdot(2\cdot n+n+n)+1$ meaning that $z\cdot z$ is odd contradicting the fact that $z\cdot z$ is even. □

6.3 Denumerability of the integers

We are going to prove now that the set of integers is **denumerable**.

Lemma 6.58. $\mathbb{N}_{0_{\mathbb{Z}}}$ *is* ***denumerable***

Proof. This follows directly from 6.22 (3) where it is proved that $i_{\mathbb{Z}}:\mathbb{N}_0\rightarrow\mathbb{N}_{0_{\mathbb{Z}}}$ is a bijection. □

Theorem 6.59. $\mathbb{Z}$ *is* ***denumerable****.*

Proof. Using 6.24 we have that $\mathbb{Z}=\mathbb{N}_{0_{\mathbb{Z}}}\bigcup(-\mathbb{N}_{0_{\mathbb{Z}}})$ where $\mathbb{N}_{0_{\mathbb{Z}}}=\{-n|n\in\mathbb{N}_{0_{\mathbb{Z}}}\}$. We can now construct the trivial bijection $f:\mathbb{N}_{0_{\mathbb{Z}}}\rightarrow(-\mathbb{N}_{0_{\mathbb{Z}}})$ defined by $n\rightarrow -n$. So we have $\mathbb{N}_{0_{\mathbb{Z}}}\approx(-\mathbb{N}_{0_{\mathbb{Z}}})$ and by the previous lemma that $\mathbb{N}_{0_{\mathbb{Z}}}\approx\mathbb{N}_0$ and thus $(-\mathbb{N}_{0_{\mathbb{Z}}})\approx\mathbb{N}_0$. This means that $\mathbb{Z}$ is the union of two denumerable sets and thus by 5.61 that $\mathbb{Z}$ is denumerable. □

Chapter 7
The rational numbers

7.1 Definition and arithmetic's

Just like we defined integer numbers using a equivalence relation on $\mathbb{N}_0$ we define the rational numbers using a equivalence relation in $\mathbb{Z}$

Definition 7.1. $\mathbb{Z}_0 = \mathbb{Z} \backslash \{0\}$ *(here* $0 = \sim[(1,1)]$*)*

Definition 7.2. $\simeq \subseteq (\mathbb{Z} \times \mathbb{Z}_0) \times (\mathbb{Z} \times \mathbb{Z}_0)$ *is the relation defined on* $\mathbb{Z} \times \mathbb{Z}_0$ *by* $\simeq = \{((n,m),(r,k)) \in \mathbb{Z} \times \mathbb{Z}_0 | n \cdot k = m \cdot \mathrm{kr}\}$*. In other words if* $(n,m),(r,k) \in \mathbb{Z} \times \mathbb{Z}_0$ *then* $(n,m) \simeq (r,k)$ *iff* $n \cdot k = m \cdot r$

Theorem 7.3. $\simeq$ *is a equivalence relation on* $\mathbb{Z} \times \mathbb{Z}_0$

Proof.

1. **(reflexive)** If $(n,m) \in \mathbb{Z} \times \mathbb{Z}_0$ then $n \cdot m \underset{6.9}{=} m \cdot n \Rightarrow (n,m) \simeq (n,m)$
2. **(symmetry)** If $(n,m),(r,k) \in \mathbb{Z} \times \mathbb{Z}_0$ and $(n,m) \simeq (r,k)$ then $n \cdot k = m \cdot r \underset{6.9}{\Rightarrow} r \cdot m = k \cdot n \Rightarrow (r,k) \simeq (n,m)$
3. **(transitive)** If $(i,j),(n,m),(r,k) \in \mathbb{Z} \times \mathbb{Z}_0$ and $(i,j) \simeq (n,m)$ and $(n, m) \simeq (r,k)$ then we have $i \cdot m = j \cdot n$ and $n \cdot k = m \cdot r$ then $(i \cdot m) \cdot k = (j \cdot n) \cdot k = j \cdot (n \cdot k) = j \cdot (m \cdot r) \Rightarrow (i \cdot k) \cdot m = (j \cdot r) \cdot m \underset{m \neq 0, 6.15}{\Rightarrow} i \cdot k = j \cdot r \Rightarrow (i,j) \simeq (k, r)$ □

Definition 7.4. *We define the set of rationals* $\mathbb{Q}$ *by* $\mathbb{Q}/\simeq$ *, we note* $\simeq[(n,k)] \in \mathbb{Q}/\simeq$ *as* $\frac{n}{k}$*, n is called the denominator, k is the nominator. We have then that* $\frac{n}{k} = \frac{n'}{k'}$ *iff* $(n,k) \sim (n',k') \Leftrightarrow n \cdot k' = k \cdot n'$

Theorem 7.5. $+ \colon \mathbb{Q} \times \mathbb{Q} \rightarrow \mathbb{Q}$ *defined by* $\left(\frac{n}{m}, \frac{r}{k}\right) \rightarrow \frac{n}{m} + \frac{r}{k} = \frac{n \cdot k + m \cdot r}{m \cdot k}$ *is a function*

Proof. First note that as $m,k \in \mathbb{Z}_0$ then $m \cdot k \in \mathbb{Z}_0$ [if $m \cdot k = 0 \underset{6.9}{\Rightarrow} m = n = 0$ contradicting $m,k \in \mathbb{Z}_0$] we have thus that $\frac{n \cdot k + m \cdot r}{m \cdot k} \in \mathbb{Q}$

Second, suppose that $\frac{n}{m} = \frac{n'}{m'}$ and $\frac{r}{k} = \frac{r'}{k'}$ then we have $n \cdot m' = m \cdot n'$ and $r \cdot k' = k \cdot r'$. Now

$$\begin{aligned}
(n \cdot k + m \cdot r) \cdot (m' \cdot k') &= (n \cdot k) \cdot (m' \cdot k') + (m \cdot r) \cdot (m' \cdot k') \\
&= (n \cdot m') \cdot (k \cdot k') + (r \cdot k') \cdot (m \cdot m') \\
&= (m \cdot n') \cdot (k \cdot k') + (k \cdot r') \cdot (m \cdot m') \\
&= (n' \cdot k') \cdot (m \cdot k) + (m' \cdot r') \cdot (m \cdot k) \\
&= (m \cdot k) \cdot (n' \cdot k' + m' \cdot r') \\
&\Rightarrow \\
\frac{n \cdot k + m \cdot r}{m \cdot k} &= \frac{n' \cdot k' + m' \cdot r'}{m' \cdot k'}
\end{aligned}$$

meaning that the function is well defined. $\square$

Theorem 7.6. *If* $k \in \mathbb{Z}_0$ *and* $\frac{a}{b} \in \mathbb{Q}$ *then* $\frac{a}{b} = \frac{a \cdot k}{b \cdot k}$

Proof. $a \cdot (b \cdot k) = (b \cdot a) \cdot k = b \cdot (a \cdot k) \Rightarrow \frac{a}{b} = \frac{a \cdot k}{b \cdot k}$ $\square$

Theorem 7.7. $\langle \mathbb{Q}, + \rangle$ *forms a* ***abelian group*** *with neutral element* $0 = \frac{0}{1}$. *Note that the same symbol* 0 *is used for neutral elements in* $\mathbb{Z}$ *and* $\mathbb{Q}$, *context tell's us always what we mean by* 0.

Proof.

1. **(associative)** If $\frac{a}{b}, \frac{c}{d}, \frac{e}{f} \in \mathbb{Q}$ then we have

$$\begin{aligned}
\frac{a}{b} + \left(\frac{c}{d} + \frac{e}{f}\right) &= \frac{a}{b} + \frac{c \cdot f + d \cdot e}{d \cdot f} \\
&= \frac{a \cdot (d \cdot f) + b \cdot (c \cdot f + d \cdot e)}{b \cdot (d \cdot f)} \\
&= \frac{(a \cdot d) \cdot f + (b \cdot c) \cdot f + (b \cdot d) \cdot e}{(b \cdot d) \cdot f} \\
&= \frac{(a \cdot d + b \cdot c) \cdot f + (b \cdot d) \cdot e}{(b \cdot d) \cdot f} \\
&= \frac{a \cdot d + b \cdot c}{b \cdot d} + \frac{e}{f} \\
&= \left(\frac{a}{b} + \frac{c}{d}\right) + \frac{e}{f}
\end{aligned}$$

2. **(neutral element)** We define the neutral element in $\mathbb{Q}$ to be $0 = \frac{0}{1}$ (note that one $0 \in \mathbb{Z}$ and the other is in $\mathbb{Q}$ context will always tell us which is which). If $\frac{a}{b} \in \mathbb{Q}$ then we have

$$\begin{aligned}
\frac{a}{b} + 0 &= \frac{a}{b} + \frac{0}{1} \\
&= \frac{a \cdot 1 + b \cdot 0}{b \cdot 1} \\
&= \frac{a + 0}{b} \\
&= \frac{a}{b} \\
0 + \frac{a}{b} &= \frac{0}{1} + \frac{a}{b} \\
&= \frac{0 \cdot b + 1 \cdot a}{1 \cdot b} \\
&= \frac{a}{b}
\end{aligned}$$

3. **(inverse)** If $\frac{a}{b} \in \mathbb{Q}$ then we prove that the inverse is $\frac{-a}{b}$.

Proof. We have

$$\begin{aligned}
\frac{a}{b}+\frac{-a}{b} &= \frac{a\cdot b+b\cdot(-a)}{b\cdot b}\\
&= \frac{b\cdot(a+(-a))}{b\cdot b}\\
&= \frac{b\cdot 0}{b\cdot b}\\
&\underset{7.6}{=} \frac{0}{b}\\
&= \frac{b\cdot 0}{b\cdot 1}\\
&= \frac{0}{1}=0\\
\frac{-a}{b}+\frac{a}{b} &= \frac{(-a)\cdot b+b\cdot a}{b\cdot b}\\
&= \frac{b\cdot(-a+a)}{b\cdot b}\\
&= \frac{b\cdot 0}{b\cdot b}\\
&= 0
\end{aligned}$$

□

4. **(commutative)** If $\frac{a}{b},\frac{c}{d}\in\mathbb{Q}$ then

$$\begin{aligned}
\frac{a}{b}+\frac{c}{d} &= \frac{a\cdot d+b\cdot c}{b\cdot d}\\
&= \frac{c\cdot b+d\cdot a}{d\cdot b}\\
&= \frac{c}{d}+\frac{a}{b}
\end{aligned}$$

□

Notation 7.8. *Whenever we write $a-b$ we mean $a+(-b)$*

Theorem 7.9. $\cdot:\mathbb{Q}\times\mathbb{Q}\rightarrow\mathbb{Q}$ *defined by* $\left(\frac{n}{m},\frac{r}{k}\right)\rightarrow\frac{n}{m}\cdot\frac{r}{k}=\frac{n\cdot r}{m\cdot k}$ *is a function.*

Proof. First as $m,k\in\mathbb{Z}_0$ we have that $m\cdot k\in\mathbb{Z}_0$ [if $m\cdot k=0\Rightarrow m=k=0$ by 6.9] so we have that $\frac{n\cdot r}{m\cdot k}\in\mathbb{Q}$. Secondly if $\frac{n}{m}=\frac{n'}{m'}$ and $\frac{r}{k}=\frac{r'}{k'}$ then we have $n\cdot m'=m\cdot n'$ and $r\cdot k'=k\cdot r'$. We have then

$$\begin{aligned}
(n\cdot r)\cdot(m'\cdot k') &= (n\cdot m')\cdot(r\cdot k')\\
&= (m\cdot n')\cdot(k\cdot r')\\
&= (m\cdot k)\cdot(n'\cdot r')\\
&\Rightarrow\\
\frac{n\cdot r}{m\cdot k} &= \frac{n'\cdot r'}{m'\cdot k'}
\end{aligned}$$

proving that the function is well-defined. □

Theorem 7.10. $\langle\mathbb{Q},+,\cdot\rangle$ *forms a* ***field****, the neutral element is* $1=\frac{1}{1}$ *and if* $\frac{a}{b}\in\mathbb{Q}\backslash\{0\}$ *the inverse for multiplication is* $\frac{b}{a}$ *(so* $\left(\frac{a}{b}\right)^{-1}=\frac{b}{a}$*). Note that* 1 *stands here for the unit in* $\mathbb{Z}$ *and* $\mathbb{Q}$*, context tell's us which is which.*

Proof. By 7.7 we have that $\langle\mathbb{Q},+\rangle$ is a abelian group, we must thus only prove the additional properties (see 3.29) of the multiplication in a field.

1. **(Distributive)** If $\frac{a}{b},\frac{c}{d},\frac{e}{f}\in\mathbb{Q}$ then we have

$$\begin{aligned}\frac{a}{b}\cdot\left(\frac{c}{d}+\frac{e}{f}\right) &= \frac{a}{b}\cdot\left(\frac{c\cdot f+e\cdot d}{d\cdot f}\right)\\ &= \frac{a\cdot(c\cdot f+e\cdot d)}{b\cdot(d\cdot f)}\\ &\underset{7.6}{=} \frac{b\cdot(a\cdot(c\cdot f+e\cdot d))}{b\cdot(b\cdot(d\cdot f))}\\ &= \frac{(a\cdot c)\cdot(b\cdot f)+(a\cdot e)\cdot(b\cdot d)}{(b\cdot d)\cdot(b\cdot f)}\\ &= \frac{a\cdot c}{b\cdot d}+\frac{a\cdot e}{b\cdot f}\\ &= \left(\frac{a}{b}\cdot\frac{c}{d}\right)+\left(\frac{a}{b}\cdot\frac{e}{f}\right)\end{aligned}$$

2. **(Neutral)** The neutral element $1=\frac{1}{1}$ (note that we have two 1's here context will always tell if $1\in\mathbb{Z}$ or $1\in\mathbb{Q}$), we have then if $\frac{a}{b}\in\mathbb{Q}$ that

$$\begin{aligned}1\cdot\frac{a}{b} &= \frac{1}{1}\cdot\frac{a}{b}\\ &= \frac{1\cdot a}{1\cdot b}\\ &= \frac{a}{b}\\ &= \frac{a\cdot 1}{b\cdot 1}\\ &= \frac{a}{b}\cdot\frac{1}{1}\\ &= \frac{a}{b}\cdot 1\end{aligned}$$

3. **(Commutative)** If $\frac{a}{b},\frac{c}{d}\in\mathbb{Q}$ then we have

$$\begin{aligned}\frac{a}{b}\cdot\frac{c}{d} &= \frac{a\cdot c}{b\cdot d}\\ &= \frac{c\cdot a}{d\cdot b}\\ &= \frac{c}{d}\cdot\frac{a}{b}\end{aligned}$$

4. **(Associative)** If $\frac{a}{b}, \frac{c}{d}, \frac{e}{f} \in \mathbb{Q}$ then we have

$$\begin{aligned}\frac{a}{b} \cdot \left(\frac{c}{d} \cdot \frac{e}{f}\right) &= \frac{a}{b} \cdot \frac{c \cdot e}{d \cdot f} \\ &= \frac{a \cdot (c \cdot e)}{b \cdot (d \cdot f)} \\ &= \frac{(a \cdot c) \cdot e}{(b \cdot d) \cdot f} \\ &= \frac{a \cdot c}{b \cdot d} \cdot \frac{e}{f} \\ &= \left(\frac{a}{b} \cdot \frac{c}{d}\right) \cdot \frac{e}{f}\end{aligned}$$

5. **(Inverse for non zero element)** If $\frac{a}{b} \in \mathbb{Q} \backslash 0$ then $\frac{a}{b} \neq \frac{0}{1}$. Now if $a = 0$ then $a \cdot 1 = 0 = b \cdot 0 \Rightarrow \frac{a}{b} = \frac{0}{1}$ a contradiction, so we must have $a \neq 0$ and thus $\frac{b}{a} \in \mathbb{Q}$. we prove now that $\left(\frac{a}{b}\right)^{-1} = \frac{b}{a}$

$$\begin{aligned}\frac{b}{a} \cdot \frac{a}{b} &= \frac{b \cdot a}{a \cdot b} \\ &= \frac{a \cdot b}{a \cdot b} \\ &= \frac{(a \cdot b) \cdot 1}{(a \cdot b) \cdot 1} \\ &\underset{7.6}{=} \frac{1}{1} \\ &= 1\end{aligned}$$

□

7.2 Power in $\mathbb{Q}$

Definition 7.11. *As $\langle \mathbb{Q}, \cdot \rangle$ is a abelian semi-group we have by 4.22 that given a $a \in \mathbb{Q}$ and $n \in \mathbb{N}_0$ that there exists a a^n such that*

$$\begin{aligned}a^0 &= 1 \\ a^{n+1} &= a^n \cdot a \underset{\text{abelian}}{=} a \cdot a^n\end{aligned}$$

Theorem 7.12. *If $n, n' \in \mathbb{N}_0$ and $a \in \mathbb{Q}$ then $a^{n'+n} = a^{n'} \cdot a^n$*

Proof. We prove this by induction on n. So let $X = \{n \in \mathbb{N}_0 | a^{n'+n} = a^{n'} \cdot a^n\}$ then we have

1. If $n = 0$ then $a^{n'+n} = a^{n'+0} = a^{n'} = a^{n'} \cdot 1 = a^{n'} \cdot a^0 \Rightarrow 0 \in X$
2. If $n \in X$ then $a^{n'+(n+1)} = a^{(n'+n)+1} = a^{(n'+n)} \cdot a \underset{n \in X}{=} (a^{n'} \cdot a^n) \cdot a = a^{n'} \cdot (a^n \cdot a) = a^{n'} \cdot a^{n+1}$ and thus $n+1 \in X$

Using mathematical induction (see 4.10) we have $X = \mathbb{N}_0$ proving the theorem $\square$

Theorem 7.13. *If* $n \in \mathbb{N}_0$ *and* $a, b \in \mathbb{Q}$ *then we have* $(a \cdot b)^n = a^n \cdot b^n$

Proof. We prove this by induction so take $\mathcal{S} = \{n \in \mathbb{N}_0 | (a \cdot b)^n = a^n \cdot b^n\}$ then we have

$\boldsymbol{n = 0}$**.** then $(a \cdot b)^0 = 1 = 1 \cdot 1 = a^0 \cdot b^0$

$\boldsymbol{n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}}$**.** then $(a \cdot b)^{n+1} = (a \cdot b)^n \cdot (a \cdot b) = (a^n \cdot b^n) \cdot (a^n \cdot a) \cdot (b^n \cdot b) = a^{n+1} \cdot b^{n+1}$ $\square$

Theorem 7.14. *In* $\mathbb{Q}$ *we have*

$$\begin{aligned} 0^n &= 0\,(\text{if } n \neq 0) \\ 1^n &= 1 \\ (-1)^n &= -1 \text{ or } 1 \\ (-1)^{2 \cdot n} &= 1 \\ (-1)^{2 \cdot n + 1} &= -1 \end{aligned}$$

Proof.

1. If $n \neq 0 \Rightarrow \exists m \in \mathbb{N}_0 \vdash n = m+1$ then $0^n = 0^{(m+1)} = 0^m \cdot 0 = 0$
2. $1^n = 1$ is proved by induction on n, let $X = \{n \in \mathbb{N}_0 | 1^n = 1\}$ then
   a. $1^0 = 1 \Rightarrow 0 \in X$
   b. If $n \in X \Rightarrow 1^{n+1} = 1^n \cdot 1 \underset{n \in X}{=} 1 \cdot 1 = 1 \Rightarrow n+1 \in X$

   so $X = \mathbb{N}_0$
3. $(-1)^n = \pm 1$ is proved by induction on n, let $X = \{n \in \mathbb{N}_0 | (-1)^n = -1 \text{ or } 1\}$ then
   a. $(-1)^0 = 1 \Rightarrow 0 \in X$
   b. If $n \in X$ then $(-1)^{n+1} = (-1)^n \cdot (-1) \underset{n \in X}{=} (-1) \cdot (-1) \vee 1 \cdot (-1) = 1 \vee -1 \Rightarrow n+1 \in X$

   so $X = \mathbb{N}_0$
4. $(-1)^{2 \cdot n} = (-1)^{(1+1) \cdot n} = (-1)^{n+n} = (-1)^n \cdot (-1)^n \underset{(3)}{=} (-1) \cdot (-1) \text{ or } 1 \cdot 1 = 1$
5. $(-1)^{2 \cdot n + 1} = (-1)^{2 \cdot n} \cdot (-1) \underset{(4)}{=} 1 \cdot (-1) = -1$ $\square$

7.3 Order relation

Theorem 7.15. $\text{sign}\colon \mathbb{Q} \to \{1, -1\} \subseteq \mathbb{N}_0$ *defined by*

$$\text{sign}\left(\frac{a}{b}\right) = \begin{cases} 1 \text{ if } 0 \leqslant a \cdot b \\ -1 \text{ if } \neg(0 \leqslant a \cdot b) \end{cases}$$

is a well defined function.

Proof. If $\frac{a}{b} = \frac{a'}{b'}$ then we must prove that $\operatorname{sign}\left(\frac{a}{b}\right) = \operatorname{sign}\left(\frac{a'}{b'}\right)$. As $\frac{a}{b} = \frac{a'}{b'}$ then $a \cdot b' = b \cdot a'$. We have now the following cases

1. $\operatorname{sign}\left(\frac{a}{b}\right) = 1 \Rightarrow 0 \leqslant a \cdot b$ then by 6.39 we have $0 \leqslant b' \cdot b'$ giving by 6.38 that $0 \cdot (b' \cdot b') \leqslant (a \cdot b) \cdot (b' b') \Rightarrow 0 \leqslant (a \cdot b') \cdot (b \cdot b') \Rightarrow 0 \leqslant (b \cdot a') \cdot (b \cdot b') \Rightarrow 0 \leqslant (a' \cdot b') \cdot (b \cdot b)$. We proceed now by contradiction that $0 \leqslant a' \cdot b'$, so assume that $a' \cdot b' < 0 \underset{6.39, 6.37}{\Rightarrow} (a' \cdot b') \cdot (b \cdot b) < 0$ contradicting $0 \leqslant (a' \cdot b') \cdot (b \cdot b)$ so $0 \leqslant a' \cdot b' \Rightarrow \operatorname{sign}\left(\frac{a'}{b'}\right) = 1$
2. $\operatorname{sign}\left(\frac{a}{b}\right) = -1 \Rightarrow \neg(0 \leqslant a \cdot b) \underset{2.144, 6.26}{\Rightarrow} a \cdot b < 0 \underset{6.39, 6.37}{\Rightarrow} (a \cdot b) \cdot (b' \cdot b') < 0 \Rightarrow (a \cdot b') \cdot (b \cdot b') < 0 \Rightarrow (a' \cdot b) \cdot (b \cdot b') < 0 \Rightarrow (a' \cdot b') \cdot (b \cdot b) < 0$. Suppose now that $0 \leqslant (a' \cdot b')$ then by 6.38,6.39 we have $0 \leqslant (a' \cdot b') \cdot (b \cdot b) < 0 \Rightarrow 0 < 0$ a contradiction, so we must have $(a' \cdot b') < 0$ or $\operatorname{sign}\left(\frac{a'}{b'}\right) = -1$ □

Lemma 7.16. *If* $\operatorname{sign}(q) = -1$ *then* $\operatorname{sign}(-q) = 1$

Proof. If $q = \frac{a}{b}$ then from $\operatorname{sign}(q) = 1$ we have $a \cdot b < 0$ and thus by 6.27 we have $-0 < -(a \cdot b) \Rightarrow 0 < (-a) \cdot b \Rightarrow 0 \leqslant (-a) \cdot b \Rightarrow \operatorname{sign}\left(\frac{-a}{b}\right) = 1 \Rightarrow \operatorname{sign}(-q) = 1$ □

Definition 7.17. (Order Relation $\mathbb{Q}$) $\leqslant \subseteq \mathbb{Q} \times \mathbb{Q}$ *is defined by* $\leqslant = \{(q, r) \in \mathbb{Q} | \operatorname{sign}(r + (-q)) = 1\}$ *(note that this definition is independent of the representation of a rational number as* sign *is a well-defined function). So we have* $q \leqslant r \Leftrightarrow \operatorname{sign}(r + (-q)) = 1$

Lemma 7.18. *If* $q \in \mathbb{Q}$ *with* $0 \leqslant q$ *and* $q \leqslant 0$ *then* $q = 0$

Proof. Let $q = \frac{a}{b}$ then $-q = \frac{-a}{b}$. From $0 \leqslant q \wedge q \leqslant 0$ we have $1 = \operatorname{sign}(q + (-0)) = \operatorname{sign}(q)$ and $1 = \operatorname{sign}(0 + (-q)) = \operatorname{sign}(-q)$ and thus $\operatorname{sign}\left(\frac{a}{b}\right) = 1 \wedge \operatorname{sign}\left(\frac{-a}{b}\right) = 1$. This gives $0 \leqslant a \cdot b \wedge 0 \leqslant (-a) \cdot b \underset{6.27}{\Rightarrow} 0 \leqslant a \cdot b \wedge (-((-a) \cdot b)) \leqslant -0 \Rightarrow 0 \leqslant a \cdot b \leqslant 0 \Rightarrow a \cdot b = 0 \underset{b \neq 0 \wedge 6.9}{\Rightarrow} a = 0 \Rightarrow q = \frac{0}{b} = \frac{0}{1} = 0$ □

Lemma 7.19. *If* $q, r, s \in \mathbb{Q}$ *with* $q \leqslant r \Rightarrow q + s \leqslant r + s$

Proof. $q \leqslant r \Rightarrow 1 = \operatorname{sign}(r + (-q)) = \operatorname{sign}(r + (s + (-s)) + (-q)) = \operatorname{sign}((r + s) + (-(q + s))) \Rightarrow q + s \leqslant r + s$ □

Lemma 7.20. *If* $q, r \in \mathbb{Q}$ *with* $0 \leqslant q$ *and* $0 \leqslant r$ *then* $0 \leqslant q + r$

Proof. Let $q = \frac{a}{b}$ and $r = \frac{a'}{b'}$. From $0 \leqslant q$ we have $1 = \operatorname{sign}(q + (-0)) = \operatorname{sign}(q)$ and from $0 \leqslant r$ we have $1 = \operatorname{sign}(r + (-0)) = \operatorname{sign}(r)$. This gives us $0 \leqslant a \cdot b$ and $0 \leqslant a' \cdot b'$. Now $\frac{a}{b} + \frac{a'}{b'} = \frac{a \cdot b' + b \cdot a'}{b \cdot b'}$ and □

Theorem 7.21. $\langle \mathbb{Q}, \leqslant \rangle$ *forms a partially ordered set that is fully-ordered*

Proof.

1. **(Reflexivity)** If $q \in \mathbb{Q} \Rightarrow q - q = 0 = \frac{0}{1} \Rightarrow \operatorname{sign}(q - q) = \operatorname{sign}\left(\frac{0}{1}\right) \underset{0 \leqslant 0 = 0 \cdot 1}{=} 1 \Rightarrow q \leqslant q$

2. **(Anti-symmetry)** If $q \leqslant r$ and $r \leqslant q$ then we have using 7.19 that $q+(-q) \leqslant r+(-q)$ and $r+(-r) \leqslant q+(-r) \Rightarrow 0 \leqslant r+(-q)$ and $0 \leqslant q+(-r) = -(r+(-q)) \Rightarrow 0 \leqslant r+(-q) \leqslant 0 \underset{7.18}{\Rightarrow} r+(-q) = 0 \Rightarrow r = q$

3. **(Transitivity)** If $q \leqslant r$ and $r \leqslant s$ then using 7.19 we have $0 = q+(-q) \leqslant r+(-q)$ and $0 = r+(-r) \leqslant s+(-r)$, using 7.20 we have then that $0 \leqslant (r+(-q)) + (s+(-r)) = s+(-q)$ which using 7.19 again gives $q = 0+q \leqslant (s+(-q))+q \Rightarrow q \leqslant s$ proving transitivity.

4. **(Fully-ordered)** Let $q, r \in \mathbb{Q}$ then for $q+(-r)$ we have either

   a. $\operatorname{sign}(q+(-r)) = 1 \Rightarrow r \leqslant q$

   b. $\operatorname{sign}(q+(-r)) = -1$ then by 7.16 we have $\operatorname{sign}(-(q+(-r))) = 1 \Rightarrow \operatorname{sign}(r+(-q)) = 1 \Rightarrow q \leqslant r$ □

Theorem 7.22. *If $q, r \in \mathbb{Q}$ with $q \leqslant r$ then we have $-r \leqslant -q$*

Proof. If $q \leqslant r$ then $\operatorname{sign}(r-q) = 1$. Now $(-q)-(-r) = (-q)+(-(-r)) = (-q)+r = r+(-q) = r-q$ and this $\operatorname{sign}((-q)-(-r)) = \operatorname{sign}(r-q) = 1 \Rightarrow -r \leqslant -q$ □

Theorem 7.23. *If $q \in \mathbb{Q} \Rightarrow q < q+1$ and $q-1 < q$*

Proof. First in $\mathbb{N}_0$ we have $0 < 1$ so that using the injectivity and order preserving off $i_{\mathbb{N}_0}$ we have $0 = \sim(s(0), 1) = i_{\mathbb{N}_0}(0) < i_{\mathbb{N}_0}(1) = \sim(s(1), 1) = 1$ so that $0 < 1$ in $\mathbb{Z}$. In the same way we have that $0 = \frac{0}{1} = i_{\mathbb{Z}}(0) < i_{\mathbb{Z}}(1) = \frac{1}{1} = 1$ so that $0 < 1$ in $\mathbb{Q}$. Now we have $0 < 1 \Rightarrow 0 \leqslant 1 \xRightarrow[7.19]{} 0+q \leqslant 1+q \Rightarrow q \leqslant q+1$ if now $q = q+1$ then $0 = q+(-q) = q+(-q)+1 = 1$ contradicting $0 < 1$ so we have that $q < q+1$. Finally $q < q+1 \Rightarrow q \leqslant q+1 \Rightarrow q+(-1) \leqslant q+1+(-1) \Rightarrow q-1 \leqslant q$, if now $q-1 = q$ then $-1 = 0 \Rightarrow 0 = 1$ a contradiction, so we have $q-1 < q$. □

Definition 7.24. $\mathbb{Z}_{\mathbb{Q}} = \left\{\frac{a}{1} | a \in \mathbb{Z}\right\}$

Theorem 7.25. *$\langle \mathbb{Z}_{\mathbb{Q}}, +, \cdot \rangle$ is a sub-ring of $\langle \mathbb{Q}, +, \cdot \rangle$ and $i_{\mathbb{Z}}: \mathbb{Z} \to \mathbb{Z}_{\mathbb{Q}}$ defined by $n \to \frac{n}{1}$ is a ring-isomorphism that is order preserving*

Proof. First we prove that $\langle \mathbb{Z}_{\mathbb{Q}}, +, \cdot \rangle$ is a sub ring of $\mathbb{Q}$. If $m, n \in \mathbb{Z}_{\mathbb{Q}}$ then there exists a $m', n' \in \mathbb{Z}$ then $m = \frac{m'}{1}, n = \frac{n'}{1}$ we have then:

1. $m+n = \frac{m'}{1} + \frac{n'}{1} = \frac{m' \cdot 1 + n' \cdot 1}{1 \cdot 1} = \frac{m'+n'}{1} \in \mathbb{Z}_{\mathbb{Q}}$
2. $-\frac{n}{1} = \frac{-n}{1} \in \mathbb{Z}_{\mathbb{Q}}$
3. $m \cdot n = \frac{m' \cdot n'}{1 \cdot 1} = \frac{m' \cdot n'}{1} \in \mathbb{Z}_{\mathbb{Q}}$
4. $1 = \frac{1}{1} \in \mathbb{Z}_{\mathbb{Q}}$
5. $0 = \frac{0}{1} \in \mathbb{Z}_{\mathbb{Q}}$

proving that $\langle \mathbb{Z}_{\mathbb{Q}}, +, \cdot \rangle$ is a sub-ring of $\langle \mathbb{Q}, +, \cdot \rangle$. Next prove that $i_{\mathbb{Z}}$ is a ring isomorphism

1. If $m, n \in \mathbb{Z}$ then $i_{\mathbb{Z}}(m+n) = \frac{m+n}{1} = \frac{m \cdot 1 + n \cdot 1}{1 \cdot 1} = \frac{m}{1} + \frac{n}{1} = i_{\mathbb{Z}}(m) + i_{\mathbb{Z}}(n)$

2. If $m, n \in \mathbb{Z}$ then $i_{\mathbb{Z}}(m \cdot n) = \frac{m \cdot n}{1} = \frac{m \cdot n}{1 \cdot 1} = \frac{m}{1} \cdot \frac{n}{1} = i_{\mathbb{Z}}(m) \cdot i_{\mathbb{Z}}(n)$
3. $i_{\mathbb{Z}}(1) = \frac{1}{1} = 1$ (note that 1 stands for the unit in $\mathbb{Z}$ and $\mathbb{Q}$)

Secondly we have to prove that $i_{\mathbb{Z}}$ is a bijection.

1. **(Injectivity)** If $i_z(m) = i_z(n) \Rightarrow \frac{m}{1} = \frac{n}{1} \Rightarrow m \cdot 1 = n \cdot 1 \Rightarrow m = n$
2. **(Surjectivity)** If $a \in \mathbb{Z}$ then $\exists a' \in \mathbb{Q}$ with $a = \frac{a'}{1} = i_{\mathbb{Z}}(a')$

Finally to prove that $i_{\mathbb{Z}}$ is order preserving, assume that $m, n \in \mathbb{Z}$ with $n \leqslant m$ then we have $i_{\mathbb{Z}}(m) + (-i_{\mathbb{Z}}(n)) = \frac{m}{1} + \frac{-n}{1} = \frac{m \cdot 1 + (-n) \cdot 1}{1 \cdot 1} = \frac{m-n}{1}$, now from $n \leqslant m$ we have $0 \leqslant m - n$ so that $(m-n) \cdot 1 = m - n \geqslant 0 \Rightarrow \text{sign}(i_{\mathbb{Z}}(m) + (-i_{\mathbb{Z}}(n))) = \text{sign}\left(\frac{m-n}{1}\right) = 1 \Rightarrow i_{\mathbb{Z}}(n) \leqslant i_{\mathbb{Z}}(m)$. □

Lemma 7.26. *If $q \in \mathbb{Q}$ with $q > 0$ then there exists $a, b \in \mathbb{N}_{0_{\mathbb{Z}}} \backslash \{0\}$ such that $q = \frac{a}{b}$*

Proof. If $q \in \mathbb{Q}$ then there exist a $a', b' \in \mathbb{Z}, b \neq 0$ such that $q = \frac{a'}{b'}$ and as $q > 0$ $a' \cdot b' \geqslant 0$. We have the following possibilities (as $a' = 0$ would mean that $q = 0$ a contradiction).

1. $a' < 0 \wedge b' > 0$ then using 6.37 $a' \cdot b' < 0 \cdot b' = 0$ contradicting $a' \cdot b' \geqslant 0$, so this case does not apply.
2. $a' > 0 \wedge b' < 0$ then using 6.37 $a' \cdot b' < 0 \cdot a' = 0$ contradicting $a' \cdot b' \geqslant 0$, so this case does not apply.
3. $a' < 0 \wedge b' < 0$ then using 6.27 we have $a = -a' > 0$, $b = -b' > 0$ and $\frac{a}{b} = \frac{(-1) \cdot a'}{(-1) \cdot b'} = \frac{a'}{b'} = q$
4. $a' > 0 \wedge b' > 0$ then we take $a = a', b = b'$ and have $q = \frac{a}{b}$ □

Lemma 7.27. *If $r \in \mathbb{Q}$ with $0 < r$ (or $0 \leqslant r$) then for $s \in \mathbb{Q}, s > 0$ we have $0 < s \cdot r$ (or $0 \leqslant s \cdot r$)*

Proof. Using the previous lemma (see 7.26) we have that $s = \frac{a}{b}$ where $a, b \in \mathbb{N}_{0_{\mathbb{Z}}} \backslash \{0\} \Rightarrow a, b > 0$. Then

1. **($0 < r$)** Again by the previous lemma we have $r = \frac{c}{d}$ where $c, d > 0$. Now $s \cdot r = \frac{a}{b} \cdot \frac{c}{d} = \frac{a \cdot c}{b \cdot d}$. Using 6.36 we have $0 < a \cdot c, 0 < b \cdot d$ and using 6.36 again we have $0 < (a \cdot c) \cdot (b \cdot d)$ so that $\text{sign}(s \cdot r - 0) = \text{sign}(s \cdot r) = 1 \Rightarrow 0 \leqslant s \cdot r$. If $s \cdot r = 0 \Rightarrow \frac{a \cdot c}{b \cdot d} = \frac{0}{1} \Rightarrow a \cdot c = 0$ contradicting $0 < a \cdot c \Rightarrow s \cdot r \neq 0 \Rightarrow 0 < s \cdot r$
2. **($0 \leqslant r$)** The we have the following cases
   a. $r = 0 \Rightarrow r \cdot s = 0 \Rightarrow 0 \leqslant r \cdot s$
   b. $0 < r \underset{(1)}{\Rightarrow} 0 < s \cdot r = < 0 \leqslant r \cdot s$

□

Theorem 7.28. *If $r, q \in \mathbb{Q}$ with $r < q$ (or $r \leqslant q$) and $s \in \mathbb{Q}, 0 < s$ then $s \cdot r < s \cdot q$ (or $s \cdot r \leqslant s \cdot q$)*

Proof.

1. **($r<q$)** In this case we have $0<(q-r)$ and by the previous lemma we have then $0<s\cdot(q-r)=s\cdot q-s\cdot r\Rightarrow s\cdot r<s\cdot q$

2. **($r\leqslant q$)** In this case $0\leqslant(q-r)$ and by the previous lemma we have then $0\leqslant s\cdot(q-r)=s\cdot q-s\cdot r\Rightarrow s\cdot r\leqslant s\cdot q$ □

Definition 7.29. $\mathbb{N}_{0_\mathbb{Q}}=\left\{\frac{a}{1}|a\in\mathbb{N}_{0_\mathbb{Z}}\right\}$

Theorem 7.30. *$\langle\mathbb{N}_{0_\mathbb{Q}},+\rangle$ is a sub-semi-group of $\langle\mathbb{Q},+\rangle$, furthermore if $a,b\in\mathbb{N}_{0_\mathbb{Q}}$ and $a\leqslant b$ then $b-a\in\mathbb{N}_{0_\mathbb{Q}}$*

Proof. Let's first prove that $\langle\mathbb{N}_{0_\mathbb{Q}},+\rangle$ is a sub-semi-group of $\langle\mathbb{Q},+\rangle$.

1. If $a,b\in\mathbb{N}_{0_\mathbb{Q}}$ then there exists $a',b'\in\mathbb{N}_{0_\mathbb{Z}}$ such that $a=\frac{a'}{1},b=\frac{b'}{1}$ and then $a+b=\frac{a'}{1}+\frac{b'}{1}=\frac{a'\cdot 1+b'\cdot 1}{1\cdot 1}=\frac{a'+b'}{1}$. As $\mathbb{N}_{0_\mathbb{Z}}$ is a semi-group (being a non empty semi-sub-group of $\mathbb{Z}$, we have $a'+b'\in\mathbb{N}_{0_\mathbb{Z}}\Rightarrow a+b\in\mathbb{N}_{0_\mathbb{Q}}$

2. $1=\frac{1}{1}\in\mathbb{N}_{0_\mathbb{Q}}$

Next if $a\leqslant b\Rightarrow\operatorname{sign}(b-a)=a$ then $b-a=b+(-a)=\frac{b'}{1}+\frac{-a'}{1}=\frac{b'-a'}{1}$ and as $\operatorname{sign}(b-a)=1$ we have then $(b'-a')\cdot 1\geqslant 0\Rightarrow b'-a'\geqslant 0\Rightarrow b'-a'\in\mathbb{N}_{0_\mathbb{Z}}\Rightarrow b-a\in\mathbb{N}_{0_\mathbb{Q}}$. □

Theorem 7.31. *$\langle\mathbb{N}_{0_\mathbb{Q}},\leqslant\rangle$ is well-ordered*

Proof. Let $A\subset\mathbb{N}_{0_\mathbb{Q}}$ be a non-empty subset of $\mathbb{Q}$ then $i_\mathbb{Z}^{-1}(A)$ is non-empty as $i_\mathbb{Z}$: $\mathbb{Z}\to\mathbb{Z}_\mathbb{Q}$ is a bijection. Also if $a\in i_\mathbb{Z}^{-1}(A)$ then $\frac{a}{1}=i_\mathbb{Z}(a)\in A\subseteq\mathbb{N}_{0_\mathbb{Q}}\Rightarrow\frac{a}{1}=\frac{a'}{1}$ where $a'\in\mathbb{N}_{0_\mathbb{Z}}\Rightarrow a\cdot 1=a'\cdot 1\Rightarrow a=a'\Rightarrow a\in\mathbb{N}_{0_\mathbb{Z}}$ so we have that $i_\mathbb{Z}^{-1}(A)\subseteq\mathbb{N}_{0_\mathbb{Z}}$. By well-ordering of $\mathbb{N}_{0_\mathbb{Z}}$ (see 6.41) there exists a $m'=\min(i_\mathbb{Z}^{-1}(A))$. Take now $m=i_\mathbb{Z}(m')$ then we have

1. $m'\in i_\mathbb{Z}^{-1}(A)\Rightarrow m=i_\mathbb{Z}(m')\in A$

2. $\forall a\in A$ we have by bijectivity of $i_\mathbb{Z}$ $a'=i_\mathbb{Z}^{-1}(a)\in i_\mathbb{Z}^{-1}(A)\Rightarrow m'\leqslant a'\Rightarrow m=i_\mathbb{Z}(m')\leqslant i_\mathbb{Z}(a')=a\Rightarrow m=\min(A)$ □

Theorem 7.32. *If $a\in\mathbb{N}_{0_\mathbb{Q}}$ then $0\leqslant a$ and if $a\neq 0$ then $1\leqslant a$*

Proof. If $a\in\mathbb{N}_{0_\mathbb{Q}}$ then there exists a $a'\in\mathbb{N}_{0_\mathbb{Z}}$ (so $0\leqslant a'$ (see 6.28)) with $a=\frac{a'}{1}$, but then $0\leqslant a'=a'\cdot 1\Rightarrow\operatorname{sign}(a-0)=\operatorname{sign}(a)=1\Rightarrow 0\leqslant a$. If $a\neq 0$ then $a'\neq 0$ and as $i_{\mathbb{N}_0}$: $\mathbb{N}_0\to\mathbb{N}_{0_\mathbb{Z}}$ is a bijection, there exists a $a''\in\mathbb{N}_0$ with $a'=i_{\mathbb{N}_0}(a'')$, as $a'\neq 0$ we must have that $a''\neq 0$ (otherwise $i_{\mathbb{N}_0}(0)=\sim[(s(0),1)]=\sim[(1,1)]=0$) and thus $0<a''$ (by 4.47). Using 4.51 we have then that $1=s(0)\leqslant a''$ and using the order preserving of $i_{\mathbb{N}_0}$ we have then $1=\sim[(s(1),1)]=i_{\mathbb{N}_0}(1)\leqslant i_{\mathbb{N}_0}(a'')=a'\Rightarrow 1\leqslant a'$. As $a-1=\frac{a'}{1}-\frac{1}{1}=\frac{a'-1}{1}$ and $(a'-1)\cdot 1=a'+(-1)\geqslant 0$ (as $a'\geqslant 1$) we have $\operatorname{sign}(a-1)=1$ or $1\leqslant a$. □

Theorem 7.33. **(Archimedean Property of $\mathbb{Q}$)** *If $x, y \in \mathbb{Q}$ with $x > 0$ then there exists a $z \in \mathbb{N}_{0_{\mathbb{Q}}}$ such that $z \cdot x > y$*

Proof. For $y \in \mathbb{Q}$ we have the following possibilities

1. $y \leqslant 0$ take then $z = 1 \in \mathbb{Q}$ then as the unit 1 in $\mathbb{Z}$ is in $\mathbb{N}_{0_{\mathbb{Z}}}$ we have that $1 = \frac{1}{1} \in \mathbb{N}_{0_{\mathbb{Q}}}$ and $z \cdot x = 1 \cdot x = x > 0 \geqslant y \Rightarrow z \cdot x > y$.
2. $0 < y$. As also $x > 0$ we have by the previous lemma (see 7.26) that there exists $p, q, r, s \in \mathbb{N}_{0_{\mathbb{Z}}} \backslash \{0\}$ such that $x = \frac{p}{q}$, $y = \frac{r}{s}$ as $p > 0, s > 0 \Rightarrow p \cdot s > 0$ then by 6.40 there exists a $z' \in \mathbb{N}_{0_{\mathbb{Z}}}$ such that $z' \cdot p \cdot s > q \cdot r \Rightarrow z' \cdot p \cdot s > q \cdot r$. Define $z = \frac{z'}{1} \in \mathbb{N}_{0_{\mathbb{Q}}}$ then $z \cdot x - y = \frac{z' \cdot p}{1 \cdot q} - \frac{r}{s} = \frac{z' \cdot p \cdot s + (-(q \cdot r))}{p \cdot q}$, as $p > 0, q > 0 \Rightarrow p \cdot q > 0 \Rightarrow (p \cdot q) \cdot (z' \cdot p \cdot s + (-(q \cdot r))) > 0 \Rightarrow \operatorname{sign}(z \cdot x - y) = 1 \Rightarrow z \cdot x > y$. □

Lemma 7.34. *If $s \in \mathbb{Q}$ then $0 < s$ iff $0 < s^{-1}$*

Proof. If $0 < s$ then by 7.26 we have $s = \frac{a}{b}$ where $a, b > 0$ as $s^{-1} = \frac{b}{a}$ we have $b \cdot a > 0 \Rightarrow \operatorname{sign}(s^{-1} - 0) = \operatorname{sign}(s^{-1}) = 1 \Rightarrow 0 < s^{-1}$. If $0 < s^{-1}$ then if $s \leqslant 0$ we have by 7.27 $s \cdot s^{-1} \leqslant s^{-1} \cdot 0 \Rightarrow 1 \leqslant 0$ a contradiction, so we must have $0 < s$. □

Lemma 7.35. *If $r, q \in \mathbb{Q}$ then we have*

1. $0 < r < 1 \Rightarrow 1 < r^{-1}$
2. $1 < r \Rightarrow r^{-1} < 1$
3. $1 < r$ *and* $0 < s \Rightarrow s < r \cdot s$
4. $r < 1$ *and* $0 < s \Rightarrow s \cdot r < s$
5. $0 < r < s \Rightarrow s^{-1} < r^{-1}$
6. $0 < r^{-1} < s^{-1} \Rightarrow s < r$
7. *If* $r \neq 0$ *then* $-(r^{-1}) = (-r)^{-1}$

Proof.

1. If $0 < r < 1 \underset{7.34}{\Rightarrow} 0 < r^{-1} \underset{7.28}{\Rightarrow} r \cdot r^{-1} < 1 \cdot r^{-1} \Rightarrow 1 < r^{-1}$
2. If $1 < r \underset{0<1}{\Rightarrow} 0 < r \underset{7.34}{\Rightarrow} 0 < r^{-1} \underset{7.28}{\Rightarrow} 1 \cdot r^{-1} < r \cdot r^{-1} \Rightarrow r^{-1} < 1$
3. If $0 < s$ and $1 < r$ then by 7.28 we have $1 \cdot s < r \cdot s = r \cdot s \Rightarrow s < r \cdot s$
4. If $0 < s$ and $r < 1$ then by 7.28 we have $r \cdot s < 1 \cdot s \Rightarrow s \cdot r < s$
5. If $0 < r < s$ then by the previous theorem we have $0 < r^{-1}, s^{-1}$ and given $r < s \underset{7.28}{\Rightarrow} r \cdot r^{-1} < s \cdot r^{-1} \Rightarrow 1 < s \cdot r^{-1} \underset{(3)}{\Rightarrow} s^{-1} < s^{-1}(s \cdot r^{-1}) = r^{-1} \Rightarrow s^{-1} < r^{-1}$
6. If $0 < r^{-1} < s^{-1}$ then using the previous lemma we have $0 < r, 0 < s$ so using $r^{-1} < s^{-1} \underset{7.28}{\Rightarrow} r^{-1} \cdot r < s^{-1} \cdot r \Rightarrow 1 < s^{-1} \cdot r \underset{(3)}{\Rightarrow} 1 \cdot s < s \cdot (s^{-1} \cdot r) = s < r$
7. If $r \neq 0$ then $-r \neq 0$ so r^{-1} and $(-r)^{-1}$ exists. Now if $r \neq 0$ then $r = \frac{a}{b}, a, b \neq 0$ and $-(r^{-1}) = -\left(\frac{b}{a}\right) = \left(\frac{-b}{a}\right) = \left(\frac{a}{-b}\right)^{-1} = \left(\frac{-1}{-1} \cdot \frac{a}{-b}\right)^{-1} = \left(\frac{-a}{b}\right)^{-1} = (-r)^{-1}$ □

Theorem 7.36. **($\mathbb{Q}$ is dense)** *If $x, y \in \mathbb{Q}$ with $x < y$ then $\exists z \in \mathbb{Q} \vdash x < z < y$*

Proof. Take $z = \frac{1}{2} \cdot (x+y) \in \mathbb{Q}$, we have then $x < y \underset{6.32}{\Rightarrow} x+x < x+y$ and $x+y < y+y$ so $\frac{2}{1} \cdot x = (1+1) \cdot x = x+x < x+y$ and $x+y < \frac{2}{1} \cdot y$. Using the fact that $0 < \frac{1}{2}$ and 6.37 we have then $x = \frac{1}{2} \cdot \frac{2}{1} \cdot x < \frac{1}{2} \cdot (x+y)$ and $\frac{1}{2} \cdot (x+y) < \frac{1}{2} \cdot \frac{2}{1} \cdot y = y$ and thus $x < \frac{1}{2} \cdot (x+y) < y$ □

The following theorem shows that $\langle \mathbb{Q}, \leqslant \rangle$ is not well-ordered and this is the reason that we extend the rational numbers to the set of real numbers.

Lemma 7.37. *$\forall r \in \mathbb{Q}$ we have $r \cdot r \neq 2 = \frac{2}{1}$*

Proof. We prove this by contradiction so assume that $\exists r' \in \mathbb{Q}$ such that $r' \cdot r' = 2$. Of course $r' \neq 0$ [otherwise $0 = r' \cdot r' = 2 \Rightarrow 0 = 2 \neq 0$ a contradiction]. If now $r' < 0$ take then $r = -r' \Rightarrow r \cdot r = (-r') \cdot (-r') = r' \cdot r' = 2$ otherwise take $r = r'$. So we have that $\exists r \in \mathbb{Q}$ with $0 < r$ and $r \cdot r = 2$. Then there exists a $m, n \in \mathbb{Z}$ such that $r = \frac{n}{m}$ and $m \neq 0$. Take now $n' = n / \gcd(n, m)$ and $m' = m / \gcd(n, m)$ then using 6.54 we have $d | n'$ and $d | m' \Rightarrow d = 1$ or $d = -1$. As $n = n' \cdot \gcd(n, m)$ and $m = m' \cdot \gcd(n, m)$ we have $r = \frac{n' \cdot \gcd(n, m)}{m' \cdot \gcd(n, m)} = \frac{n'}{m'}$. Now if $r \cdot r = 2 \Rightarrow \frac{n' \cdot n'}{m' \cdot m'} = \frac{2}{1} \Rightarrow n' \cdot n' = 2 \cdot m' \cdot m'$ then we have that $n' \cdot n'$ is even and using 6.57 we have that n' is even and thus there exists a $k \in \mathbb{Z}$ such that $n' = 2 \cdot k$. From this it follows that $2 \cdot m' \cdot m' = n' \cdot n' = 2 \cdot k \cdot 2 \cdot k \Rightarrow m' \cdot m' = 2 \cdot (k \cdot k)$ proving that $m' \cdot m'$ is even and by 6.57 that m' is even and thus the existence of a $l \in \mathbb{Z}$ such that $m' = 2 \cdot l$. We have thus that $2 | m'$ and $2 | n'$ but then we must have either $2 = -1$ or $2 = 1$ both of which are impossible so we reach a contradiction. □

Theorem 7.38. *$\langle \mathbb{Q}, \leqslant \rangle$ is not conditional complete, so there exists a non-empty set which is bounded above that does not have a least upper bound.*

Proof. Consider the set $A = \left\{ r \in \mathbb{Q} \middle| r > 0 \wedge r \cdot r < \frac{2}{1} \right\}$, then as $0 < \frac{4}{3}$ and $\frac{2}{1} - \frac{4}{3} \cdot \frac{4}{3} = \frac{18 - 16}{9} = \frac{2}{9} > 0 \Rightarrow \frac{4}{3} \cdot \frac{4}{3} < 2$ we have $\frac{4}{3} \in A$ so that $A \neq \emptyset$. Also if $x \in A$ then $x \leqslant \frac{2}{1}$ [if $\frac{2}{1} < x \underset{x \in A \Rightarrow 0 < x}{\Rightarrow} \frac{2}{1} \cdot x < x \cdot x$ and $\frac{2}{1} \cdot \frac{2}{1} < \frac{2}{1} \cdot x \Rightarrow \frac{4}{1} < x \cdot x \Rightarrow \frac{2}{1} < x \cdot x$ contradicting $x \in A \Rightarrow x \cdot x < \frac{2}{1}$] so $\frac{2}{1}$ is a upper bound of A. We prove now by contradiction that $u = \sup(A)$ does not exists . So assume that $\sup(A)$ exists. As $\frac{4}{2} \in A$ and $\frac{4}{3} - 1 = \frac{4 - 3}{3} = \frac{1}{3} > 0 \Rightarrow 1 < \frac{4}{3} \leqslant u \Rightarrow 0 < 1 < u$ we can have now only the following cases for $u \cdot u$ ($\langle \mathbb{Q}, \leqslant \rangle$ is fully-ordered)

1. $\boldsymbol{(u \cdot u = \frac{2}{1})}$ Using the previous lemma we have shown that this is impossible.

2. $\boldsymbol{(u \cdot u < \frac{2}{1})}$ Then $u \in A$. Now given a $n \in \mathbb{N}_{0_{\mathbb{Z}}} \setminus \{0\}$ we have

$$\begin{aligned}\left(u + \frac{1}{n}\right)\left(u + \frac{1}{n}\right) &= u \cdot u + u \cdot \frac{1}{n} + u \cdot \frac{1}{n} + \frac{1}{n} \cdot \frac{1}{n} \\ &= u \cdot u + 2 \cdot u \cdot \frac{1}{n} + \frac{1}{n} \cdot \frac{1}{n}\end{aligned}$$

and thus

$$\left(u+\frac{1}{n}\right)\cdot\left(u+\frac{1}{n}\right)<\frac{2}{1} \Leftrightarrow u\cdot u+2\cdot u\cdot\frac{1}{n}+\frac{1}{n}\cdot\frac{1}{n}<\frac{2}{1}$$
$$\Leftrightarrow 2\cdot u\cdot\frac{1}{n}+\frac{1}{n}\cdot\frac{1}{n}<\frac{2}{1}-u\cdot u$$

Since $\frac{2}{1}$ is a upper bound of A we have $u\leqslant\frac{2}{1}$. Now if $n\in\mathbb{N}_{0_{\mathbb{Z}}}\backslash\{0\}$ then $0<n$ and using 6.46 we have then $1\leqslant n\Rightarrow 0\leqslant n-1$ and as $\frac{1}{n}-\frac{1}{n}\cdot\frac{1}{n}=\frac{1}{n}-\frac{1}{n\cdot n}=\frac{n-1}{n\cdot n}\geqslant 0\Rightarrow\frac{1}{n}\cdot\frac{1}{n}\leqslant\frac{1}{n}$ and so we have

$$\begin{aligned}2\cdot u\cdot\frac{1}{n}+\frac{1}{n}\cdot\frac{1}{n} &\leqslant \frac{2}{1}\cdot\frac{2}{1}\cdot\frac{1}{n}+\frac{1}{n}\\ &\leqslant \frac{5}{n}\end{aligned}$$

Since $0<5$ and by assumption $0<2-u\cdot u$ we have by the Archimedean property (see 7.33) the existence of a $n_0'\in\mathbb{N}_{0_{\mathbb{Q}}}$ (thus $n_0'=\frac{n_0}{1}, n_0\in\mathbb{N}_{0_{\mathbb{Z}}}$) such that

$$5 < n_0'\cdot(2-u\cdot u)=\frac{n_0}{1}\cdot(2-u\cdot u)$$

as $n_0\neq 0$ (otherwise $5=0$) we have then by multiplying both sides of the above by $\frac{1}{n_0}$ that

$$\frac{5}{n_0} < 2-u\cdot u$$

and this gives that $2\cdot u\cdot\frac{1}{n_0}+\frac{1}{n_0}\cdot\frac{1}{n_0}\leqslant\frac{5}{n_0}<\frac{2}{1}-u\cdot u$ and thus that $\left(u+\frac{1}{n_0}\right)\cdot\left(u+\frac{1}{n_0}\right)<\frac{2}{1}$ which as $0<u<u+\frac{1}{n_0}$ means that $u+\frac{1}{n_0}\in A$ contradicting then fact that u is a upper-bound of A. Hence this case is impossible.

3. $\left(\frac{\mathbf{2}}{\mathbf{1}}<\boldsymbol{u}\cdot\boldsymbol{u}\right)$ If $n\in\mathbb{N}_{0_{\mathbb{Z}}}\backslash\{0\}$ then $u-\frac{1}{n}$ can not be a upper bound of A as $u-\frac{1}{n}<u$ (because $0<\frac{1}{n}$) and u is the least upper bound. So there exists a $r\in A$ such that $u-\frac{1}{n}<r$. Now as $0<n$ we have by 6.46 $1\leqslant n\Rightarrow 0\leqslant n-1$ and thus $1-\frac{1}{n}=\frac{n-1}{n}\geqslant 0\Rightarrow\frac{1}{n}\leqslant 1<u\Rightarrow 0<u-\frac{1}{n}$ and thus from $u-\frac{1}{n}<r$ we have $\left(u-\frac{1}{n}\right)\cdot\left(u-\frac{1}{n}\right)<\left(u-\frac{1}{n}\right)\cdot r$ and $\left(u-\frac{1}{n}\right)\cdot r<r\cdot r\Rightarrow\left(u-\frac{1}{n}\right)\cdot\left(u-\frac{1}{n}\right)<r\cdot r<\frac{2}{1}\Rightarrow\left(u-\frac{1}{n}\right)\cdot\left(u-\frac{1}{n}\right)<\frac{2}{1}\Rightarrow u\cdot u-\frac{2}{1}\cdot u\cdot\frac{1}{n}+\frac{1}{n}\cdot\frac{1}{n}<\frac{2}{1}$ and thus we have $u\cdot u-\frac{2}{1}\cdot u\cdot\frac{1}{n}<u\cdot u-\frac{2}{1}\cdot u\cdot\frac{1}{n}+\frac{1}{n}\cdot\frac{1}{n}<\frac{2}{1}$. Now $0<\frac{2}{1}\cdot u$ and $0<u\cdot u-\frac{2}{1}$ so using the Archimedean property there exists a $n_0\in\mathbb{N}_{0_{\mathbb{Z}}}$ such that $\frac{2}{1}\cdot u<n_0\cdot\left(u\cdot u-\frac{2}{1}\right)$ and as $n_0\neq 0$ [otherwise $2\cdot u<0$] we have $n\in\mathbb{N}_{0_{\mathbb{Z}}}\backslash\{0\}$ but also by multiplying the last expression by $\frac{1}{n_0}$ that $\frac{2}{n_0}\cdot u<u\cdot u-\frac{2}{1}\Rightarrow\frac{2}{1}<u\cdot u-\frac{2}{n_0}\cdot u=u\cdot u-\frac{2}{1}\cdot u\cdot\frac{1}{n_0}<2\Rightarrow 2<2$ a contradiction. So this case is also impossible.

As (1),(2) and (3) can never occur we have finally reach a contradiction. So $\sup(A)$ does not exists. □

7.4 Denumerability of the rationals

Lemma 7.39. *The sets $\mathbb{N}_{0_{\mathbb{Q}}}$ and $\mathbb{Z}_{\mathbb{Q}}$ are denumerable*

Proof. The proof that $\mathbb{Z}_{\mathbb{Q}}$ is denumerable is simple, using 7.25 we have that $\mathbb{Z}_{\mathbb{Q}} \approx \mathbb{Z}$ and as we know by 6.59 that $Z \approx \mathbb{N}_0$ we have that $\mathbb{Z}_{\mathbb{Q}} \approx \mathbb{N}_0$, so we have that $\mathbb{Z}_{\mathbb{Q}}$ is enumerable. To prove that $\mathbb{N}_{0_{\mathbb{Q}}}$ is denumerable first note that by definition (see 7.29) we have $\mathbb{N}_{0_{\mathbb{Q}}} = \left\{\frac{a}{1} | a \in \mathbb{N}_{0_{\mathbb{Z}}}\right\}$. Define now $f: \mathbb{N}_{0_{\mathbb{Z}}} \rightarrow \mathbb{N}_{0_{\mathbb{Q}}}$, we prove then that f is bijective:

1. **(injective)** If $f(a) = f(a') \Rightarrow \frac{a}{1} = \frac{a'}{1} \Rightarrow a \cdot 1 = a' \cdot 1 \Rightarrow a = a'$
2. **(surjective)** If $y \in \mathbb{N}_{0_{\mathbb{Q}}} \Rightarrow \exists a \in \mathbb{N}_{0_{\mathbb{Z}}}$ such that $y = \frac{a}{1} = f(a)$

So we have that $\mathbb{N}_{0_{\mathbb{Z}}} \approx \mathbb{N}_{0_{\mathbb{Q}}}$, as by 6.58 we have $\mathbb{N}_{0_{\mathbb{Z}}} \approx \mathbb{N}_0$ we must have $\mathbb{N}_{0_{\mathbb{Q}}} \approx \mathbb{N}_0$ and thus that $\mathbb{N}_{0_{\mathbb{Q}}}$ is denumerable. $\square$

Theorem 7.40. *The set $\mathbb{Q}$ is denumerable*

Proof. Define the mapping $f: \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Q}$ by

$$f(x, y) = \begin{cases} \frac{x}{y} \text{ if } (x, y) \in \mathbb{Z} \times \mathbb{Z}_0 = \mathbb{Z} \times (\mathbb{Z} \backslash \{0\}) \\ 0 \text{ if } (x, y) \in \mathbb{Z} \times \{0\} \end{cases}$$

then we prove that f is a surjection. If $y \in \mathbb{Q}$ then there exists a $a \in \mathbb{Z}, b \in \mathbb{Z}_0$ such that $y = \frac{a}{b} = f(a, b) \Rightarrow y = f(a, b)$. Now by 6.59 we have that $\mathbb{Z}$ is denumerable and thus by 5.59 we have that $\mathbb{Z} \times \mathbb{Z}$ is denumerable so there exists a bijection g: $\mathbb{N}_0 \rightarrow \mathbb{Z} \times \mathbb{Z}$ and thus a surjection $f \circ g: \mathbb{N}_0 \rightarrow \mathbb{Q}$. Using 5.67 we have then that $\mathbb{Q}$ is countable and thus finite or denumerable. As $\mathbb{Q}$ contains the denumerable and thus infinite subset we have by 5.36 that $\mathbb{Q}$ is infinite and thus denumerable. $\square$

Chapter 8

The real numbers

8.1 Definition

Definition 8.1. (Dedekind's Cut) *A subset $\alpha \subseteq \mathbb{Q}$ is a Dedekind's cut if the following properties are true*

1. $\alpha \neq \emptyset$
2. $\alpha \neq \mathbb{Q}$
3. $\forall r \in \alpha \wedge \forall s \in \mathbb{Q} \backslash \alpha$ *we have* $r < s$
4. α *does not have a greatest element*

Definition 8.2. ($\mathbb{R}$) *We define the set of real numbers to be the set of Dedekind's cuts.* $\mathbb{R} = \{\alpha \subseteq \mathbb{Q} | \alpha \,\mathrm{is\,a\,Dedekind's\,cut}\}$

Lemma 8.3. $\forall \alpha \in \mathbb{R}$, $\forall r \in \alpha$ *and* $\forall s \in \mathbb{Q} \vdash s \leqslant r$ *we have* $s \in \alpha$

Proof. We prove this by contradiction, so assume that there exists a $\alpha \in \mathbb{R}$ and a $s \in \mathbb{Q}$ with $s \leqslant r$ and $s \notin \alpha \Rightarrow s \in \mathbb{Q} \backslash \alpha \underset{8.1,3}{\Rightarrow} r < s \leqslant r \Rightarrow r < r$ a contradiction. □

Theorem 8.4. (Rational cuts) *If $r \in \mathbb{Q}$ then $\alpha_r = \{x \in \mathbb{Q} | x < r\}$ is a cut, called a rational cut. Furthermore we have*

1. $\alpha_r = \alpha_s$ *iff* $r = s$
2. α *is a rational cut iff* $r = \min(\mathbb{Q} \backslash \alpha)$ *exists and in that case* $\alpha = \alpha_r$

Proof.

First we prove that α_r is a cut.

1. If $r \in \mathbb{Q}$ then there exists $a, b \in \mathbb{Z}$ with $b \neq 0$ such that $r = \frac{a}{b}$, now if $b < 0$ then we can always take $a' = -a, b' = -a$ such that $\frac{a'}{b'} = \frac{-a}{-b} = \frac{(-1) \cdot a}{(-1) \cdot b} = \frac{a}{b} = q$. So we can always assume that $b > 0$. Taken now $q = \frac{a+(-1)}{b}$ then $r - q = \frac{a}{b} + \frac{(a+(-1))}{b} = \frac{-1}{b}$ and as $0 < b$ we have $-b < 0 \Rightarrow (-1) \cdot b < 0 \Rightarrow q < r$ ($q = r$ is impossible as $b \cdot (a+(-1)) \cdot b = b \cdot a + (-b) \neq b \cdot a$). This proves that $\alpha_r \neq \emptyset$.

2. As we don't have $r < r$ we have that $r \notin \alpha_r \Rightarrow \alpha_r \neq \mathbb{Q}$

3. If $w \in \alpha_r$ then $w < r$ and if $u \in \mathbb{Q} \backslash \alpha_r$ then $\neg(u < r) \Rightarrow r \leqslant u$ and thus $w < u$

4. If $m = \max(\alpha_r)$ is the greatest element of α_r then as $m \in \alpha_r$ we have $m < r$ then by the dense theorem of rational numbers (see 7.36) there exists a $q \in \mathbb{Q}$ such that $m < q < r \Rightarrow q \in \alpha_r$ contradiction the definition of a greatest element.

Next we prove that $\alpha_r = \alpha_s \Leftrightarrow r = s$

1. $\Rightarrow$ If $r \neq s$ then we have either

   a. $r < s \Rightarrow r \in \alpha_s \backslash \alpha_r \Rightarrow \alpha_r \neq \alpha_s$ a contradiction

   b. $s < r \Rightarrow s \in \alpha_r \backslash \alpha_s \Rightarrow \alpha_s \neq \alpha_r$ a contradiction

   So we conclude that $r = s$

2. $\Leftarrow$ If $r = s$ then clearly $\alpha_r = \alpha_s$

Finally we prove the equivalence of 'α is a rational cut iff the least element r of $\mathbb{Q} \backslash \alpha$ exists and in that case $\alpha = \alpha_r$'.

1. If α is a rational cut then $\exists r \in \mathbb{Q} \vdash \alpha = \alpha_r$. Then as $r \in \alpha_r = \alpha$ we have that $r \in \mathbb{Q} \backslash \alpha$. If $s \in \mathbb{Q} \backslash \alpha$ then $s \notin \alpha = \alpha_r \Rightarrow \neg(s < r)$ which by the fact that $\langle \mathbb{Q}, \leqslant \rangle$ is fully-ordered means that $r \leqslant s$ and thus that r is the least element of $\mathbb{Q} \backslash \alpha$ or $r = \min(\mathbb{Q} \backslash \alpha)$.

2. Assume that α is a cut and that $r = \min(\mathbb{Q} \backslash \alpha)$ exists then we claim that $\alpha = \alpha_r$.

   a. If $s \in \alpha$ then as $r \in \mathbb{Q} \backslash \alpha$ we have by the definition of a cut (see 8.1-3) that $s < r \Rightarrow s \in \alpha_r$. So we have $\alpha \subseteq \alpha_r$

   b. If $s \in \alpha_r \subseteq \mathbb{Q}$ then we have $s < r$ and thus $s \notin \mathbb{Q} \backslash \alpha$ (otherwise r would not be the least element). From this it follows that $s \in \alpha \Rightarrow \alpha_r \subseteq \alpha$.

   (a) and (b) gives finally $\alpha = \alpha_r$ $\square$

A consequence of the above and the fact that $\mathbb{Q}$ is not empty proves the following corollary.

Corollary 8.5. $\mathbb{R} \neq \emptyset$ *The set of real numbers is not empty.*

Definition 8.6. *We define the set* $\mathbb{Q}_{\mathbb{R}}$ *as follows* $\mathbb{Q}_{\mathbb{R}} = \{\alpha_r | r \in \mathbb{Q}\}$, *so* $\mathbb{Q}_r$ *is the set of all the rational cuts.*

Theorem 8.7. *If* $\alpha \in \mathbb{R}$ *then* $\forall \varepsilon \in \mathbb{Q} \backslash \{0\}$ *there* $\exists r \in \alpha$ *and* $\exists s \notin \alpha$ *such that* $s - r < \varepsilon$

Proof. Take $\alpha \in \mathbb{R}$ and $\varepsilon \in \mathbb{Q}$ by the definition of a cut (see 8.1 1,2) there exists a $r' \in \alpha$ and a $s' \notin \alpha$. Using the definition again (see 8.1 3) we have $r' < s'$. Using the Archimedean property of $\mathbb{Q}$ (see 7.33) there exists a $k \in \mathbb{N}_{0_{\mathbb{Q}}}$ such that $s' - r' < k \cdot \varepsilon$. Now as $r' < s' \Rightarrow 0 < s' - r'$ we can't have that $k = 0 \Rightarrow 0 < k$ so n^{-1} exists, also if $k^{-1} \leqslant 0 \underset{0<k}{\Rightarrow} k \cdot k^{-1} \leqslant 0 \Rightarrow 1 \leqslant 0 < 1$ a contradiction, so we must conclude that $0 < k^{-1}$. From $s' - r' < k \cdot \varepsilon$ it follows then that $k^{-1} \cdot (s' - r') < k^{-1} \cdot k \cdot \varepsilon \Rightarrow k^{-1} \cdot (s' - r') < \varepsilon$. Define now $A = \{n \in \mathbb{N}_{0_{\mathbb{Q}}} | r' + (n \cdot k^{-1}) \cdot (s' - r') \notin \alpha\} \subseteq \mathbb{N}_{0_{\mathbb{Q}}}$. As we have $r' + (k \cdot k^{-1}) \cdot (s' - r') = r' + (s' - r') = s' \notin \alpha$ we have that $k \in A \Rightarrow A$ is non empty. Using the fact that $\mathbb{N}_{0_{\mathbb{Q}}}$ is well-ordered (see 7.31) there exists a $k' = \min(A)$. Now if $k' = 0$ then we would have that $r' + 0 \cdot (s' - r') = r' \notin \alpha$ contradicting that $r' \in \alpha$ so we can't have $k' = 0$, using 7.32 we must then have $k' \geqslant 1$. Using 7.30 we have then that $k' - 1 \in \mathbb{N}_{0_{\mathbb{Q}}}$. By the definition of a minimum and $k' - 1 < k'$ (see 7.23) we have also $k' - 1 \in A \underset{k'-1 \in \mathbb{N}_{0_{\mathbb{Q}}}}{\Rightarrow} r' + (k' - 1) \cdot k^{-1} \cdot (s' - r') \in \alpha$. If we now define

$$
\begin{aligned}
r &= r' + ((k' - 1) \cdot k^{-1}) \cdot (s' - r') \\
s &= r' + (k' \cdot k^{-1}) \cdot (s' - r')
\end{aligned}
$$

then we have that $r \in a$ and $s \notin \alpha$ (as $k' \in A$). Now $s - r = r' + (k' \cdot k^{-1}) \cdot (s' - r') - (r' + ((k' - 1) \cdot k^{-1}) \cdot (s' - r')) = (s' - r') \cdot k^{-1} < \varepsilon$. □

Theorem 8.8. (Negative cut) *If* $\alpha \in \mathbb{R}$ *then* $-\alpha = \{-s | s \notin \alpha \wedge s \neq \min(\mathbb{Q} \backslash \alpha)\}$ *is a Dedekind's cut called the negative cut of* α*. Note that* $s \neq \min(\mathbb{Q} \backslash \alpha)$ *does not mean that* $\min(\mathbb{Q} \backslash \alpha)$ *must exists, but that* α *may not be the minimum element of* $\mathbb{Q} \backslash \alpha$ *(in other words if the minimum exists it may not be* α*).*

Proof.

1. Since $\alpha \neq \mathbb{Q}$ there exists a $\exists s \in \mathbb{Q} \vdash s \notin \alpha$ then $s + 1 \notin \alpha$ [if $s + 1 \in \alpha$ then as $s < s + 1$ (see 7.23) and $s \in \mathbb{Q} \backslash \alpha$ we have by 8.1 that $s + 1 < s < s + 1 \Rightarrow s + 1 < s + 1$ a contradiction]. As $s, s + 1 \in \mathbb{Q} \backslash \alpha$, $s < s + 1$ we can't have $s + 1 = \min(\mathbb{Q} \backslash \alpha) \Rightarrow s + 1 \in -\alpha \Rightarrow -\alpha \neq \emptyset$

2. Since $\alpha \neq \emptyset$ there exists a $r \in \alpha \subseteq \mathbb{Q}$ then $-r \notin \alpha \Rightarrow \alpha \neq \mathbb{Q}$

3. Let now $r \in -\alpha$ and $s \in \mathbb{Q} \backslash (-a)$ suppose now that $s \leqslant r \underset{7.22}{\Rightarrow} -r \leqslant -s$. From $r \in -\alpha$ we have $-r \notin \alpha$ as $s \in \mathbb{Q} \backslash (-\alpha)$ we have $s \notin -\alpha$ and thus $\neg(-s \in \alpha \wedge -s \neq \min(\mathbb{Q} \backslash \alpha)) \Rightarrow -s \notin \alpha \vee -s = \min(\mathbb{Q} \backslash \alpha)$ so we have the following cases to consider

   a. $-s \in \alpha$. Now $-r \leqslant -s \Rightarrow \neg(-s < -r)$ together with $-s \in \alpha$ and $-r \in \mathbb{Q} \backslash \alpha$ contradicts the definition of α as a cut (see 8.1 , 3). So this cases is impossible.

   b. $-s = \min(\mathbb{Q} \backslash \alpha)$ As $-s \leqslant -r$ we must have $-s = -r \Rightarrow s = r \Rightarrow s \in (-\alpha) \bigcap (\mathbb{Q} \backslash (-\alpha)) = \emptyset$ a contradiction. So this case is also impossible.

   As (a) and (b) are impossible we must conclude that $r < s$

4. We prove by contradiction that $\max(-\alpha)$ does not exists. So assume that $\max(-\alpha)$ exists, then as $\max(-\alpha) \in -\alpha$ we must have that $-\max(-\alpha) \notin \alpha \Rightarrow -\max(-\alpha) \in \mathbb{Q} \backslash \alpha$ and $-\max(-\alpha) \neq \min(\mathbb{Q} \backslash \alpha)$. We consider now the two only cases for $\min(\mathbb{Q} \backslash \alpha)$

   a. $\min(\mathbb{Q} \backslash \alpha)$ does not exists. If now $\forall s \in \mathbb{Q} \backslash \alpha$ we have $-\max(-\alpha) \leqslant s$ we would have as $-\max(-\alpha) \in \mathbb{Q} \backslash \alpha$ that $-\max(-\alpha) = \min(\mathbb{Q} \backslash \alpha)$ contradicting the fact that we assumed that $\min(\mathbb{Q} \backslash \alpha)$ does not exists. So we must have that $\exists s \in \mathbb{Q} \backslash \alpha$ so that $s < -\max(-\alpha) \Rightarrow \max(-\alpha) < -s$, as $s \in \mathbb{Q} \backslash \alpha$ and $\min(\mathbb{Q} \backslash \alpha)$ does not exists so $-s \neq \min(\mathbb{Q} \backslash \alpha)$ we must have that $-s \in -\alpha \Rightarrow -s \leqslant \max(-\alpha)$ but this contradicts $\max(-\alpha) < -s$. So we reach a final contradiction.

   b. $\min(\mathbb{Q} \backslash \alpha)$ exists. As $-\max(-\alpha) \in \mathbb{Q} \backslash \alpha$ and $-\max(-\alpha) \neq \min(\mathbb{Q} \backslash \alpha)$ we have $\min(\mathbb{Q} \backslash \alpha) < \max(-\alpha)$. Using the density theorem (see 7.36) there exists a $s \in \mathbb{Q}$ such that $\min(\mathbb{Q} \backslash \alpha) < s < -\max(-\alpha)$. As $\min(\mathbb{Q} \backslash \alpha) \in \mathbb{Q} \backslash \alpha$ we would have by the fact that α is a cut (see 8.1 3) that we can not have $s \in \alpha$ [as this would mean [$\min(\mathbb{Q} \backslash \alpha) \in \mathbb{Q} \backslash \alpha$] $s < \min(\mathbb{Q} \backslash \alpha) < s$ a contradiction]. So $s \in \mathbb{Q} \backslash \alpha$ and $\min(\mathbb{Q} \backslash \alpha) \neq s \Rightarrow -s \in -\alpha$, from $s < -\max(-\alpha) \underset{7.22}{\Rightarrow} -s \leqslant \max(-\alpha) < -s \Rightarrow -s < -s$ which is a contradiction.

   As all the possible cases gives a contradiction we must conclude that $\max(-\alpha)$ does not exists.

(1),(2),(3) and (4) proves that indeed $-\alpha$ is a cut. $\square$

For rational cuts there is a easy way to construct its negative cut as is expressed in the following theorem.

Theorem 8.9. *If* $r \in \mathbb{Q}$ *then* $-\alpha_r = \alpha_{-r}$

Proof.

$$\begin{array}{rcl} x \in -\alpha_r & \Leftrightarrow & -x \notin \alpha_r \wedge -x \neq \min(\mathbb{Q}\backslash\alpha_r) \\ & \underset{8.4 \Rightarrow r = \min(\mathbb{Q}\backslash\alpha_r)}{\Leftrightarrow} & -x \notin \alpha_r \wedge -x \neq r \\ & \Leftrightarrow & r \leqslant -x \wedge -x \neq r \\ & \Leftrightarrow & -x > r \\ & \Leftrightarrow & x < -r \\ & \Leftrightarrow & x \in \alpha_{-r} \\ & & \square \end{array}$$

8.2 Arithmetic's on $\mathbb{R}$

8.2.1 Addition in $\mathbb{R}$

Definition 8.10. *If $\alpha, \beta \in \mathbb{R}$ then we define $\alpha+\beta = \{r+s | r \in \alpha, s \in \beta\}$*

Lemma 8.11. *$\forall \alpha \in \mathbb{R}$ and $\varepsilon \in \mathbb{Q}\backslash\{0\}$ there exists a $r \in \alpha$ such that $r+\varepsilon \notin \alpha$*

Proof. Using 8.7 we have there exists a $r \in \alpha$ and a $s \notin \alpha$ such that $s-r < \varepsilon \Rightarrow s < r+\varepsilon$. Assume now that $r+\varepsilon \in \alpha$ then as $s \notin \alpha \Rightarrow s \in \mathbb{Q}\backslash\alpha$ we have by 8.1, 3 that $r+\varepsilon < s$ contradicting $s < r+\varepsilon$ so we must have that $r+\varepsilon \notin \alpha$ $\square$

Theorem 8.12. *$\forall \alpha, \beta \in \mathbb{R}$ then $\alpha+\beta \in \mathbb{R}$*

Proof. Given $\alpha, \beta \in \mathbb{R}$ we must prove that $\alpha+\beta$ is a Dedekind's cut.

1. **($\boldsymbol{\alpha+\beta \neq \emptyset}$)** Since $\alpha \neq \emptyset \wedge \beta \neq \emptyset \Rightarrow \exists r \in \alpha \wedge \exists s \in \beta \Rightarrow r+s \in \alpha+\beta \Rightarrow \alpha+\beta \neq \emptyset$
2. **($\boldsymbol{\alpha+\beta \neq \mathbb{Q}}$)** Given $\varepsilon = \frac{1}{2} \in \mathbb{Q}$ there exists by 8.11 a $r' \in \alpha$ and a $s' \in \beta$ such that $r'+\frac{1}{2} \notin \alpha$ and $s'+\frac{1}{2} \in \beta$. Now $\forall r \in \alpha$ and $\forall s \in \beta$ we have by 8.1 , 3 that $r < r'+\frac{1}{2}$ and $s < s'+\frac{1}{2} \Rightarrow r+s < r'+s'+\frac{1}{1+1} + \frac{1}{1+1} = r'+s'+1 \Rightarrow r+s < r'+s'+1$. Now if $r'+s'+1 \in \alpha+\beta$ then there exists a $r \in \alpha$ and a $s \in \beta$ such that $r'+s'+1 = r+s$ and by the above we would have the contradiction $r'+s'+1 = r+s < r'+s'+1$ so we conclude that $r'+s'+1 \notin \alpha+\beta \Rightarrow \alpha+\beta \neq \mathbb{Q}$
3. **($\boldsymbol{\forall w \in \alpha+\beta \wedge \forall v \in \mathbb{Q}\backslash(\alpha+\beta)}$ we have $\boldsymbol{w < v}$)** Let $w \in \alpha+\beta$ then there exists a $r \in \alpha$ and a $s \in \beta$ such that $w = r+s$. Assume now that there exists a $v \in \mathbb{Q}\backslash(\alpha+\beta)$ such that $v \leqslant w$. We have then $v \leqslant r+s \Rightarrow v-s \leqslant r$, using 8.3 we have then $v-s \in \alpha$ and thus $v = (v-s)+s \in \alpha+\beta$ contradicting $v \in \mathbb{Q}\backslash(\alpha+\beta)$ and thus our assumption. So we must have $\forall v \in \mathbb{Q}\backslash(\alpha+\beta)$ that $w < v$.
4. **($\boldsymbol{\alpha+\beta}$ does not have $\boldsymbol{a}$ maximum)** Suppose that $m = \max(\alpha+\beta)$ then as $m \in \alpha+\beta$ there exists a $r \in \alpha$ and a $\beta \in \beta$ such that $m = r+s$. As $\max(\alpha)$ and $\max(\beta)$ does not exists there exists a $r' \in \alpha$ such that $r < r' \Rightarrow r+s < r'+s$ which as $r'+s \in \alpha+\beta$ would give $r'+s \leqslant m = r+s < r'+s \Rightarrow r'+s < r'+s$ a contradiction. The only conclusion left is that $\alpha+\beta$ does not have a maximum

. $\square$

Theorem 8.13. *$\langle \mathbb{R}, + \rangle$ forms a abelian group, the identity element is α_0 which we will note as usually as 0 (so $0 = \alpha_0$ where the 0 in the lower index is the identity in $\langle \mathbb{Q}, + \rangle$). If $\alpha \in \mathbb{R}$ then its inverse is $-\alpha$.*

Proof. We rely here heavily on the fact that $\langle \mathbb{Q}, + \rangle$ forms a abelian group (see 7.7)

1. **(associative)** If $\alpha, \beta, \gamma \in \mathbb{R}$ then we have

$$\begin{aligned} z \in (\alpha+\beta)+\gamma &\Leftrightarrow z = r+s \wedge r \in (\alpha+\beta) \wedge s \in \gamma \\ &\Leftrightarrow z = (u+t)+s \wedge u \in \alpha \wedge t \in \beta \wedge s \in \gamma \\ &\underset{7.7}{\Leftrightarrow} z = u+(t+s) \wedge u \in \alpha \wedge t \in \beta \wedge s \in \gamma \\ &\Leftrightarrow z = u+v \wedge u \in \alpha \wedge v \in \beta+\gamma \\ &\Leftrightarrow z \in \alpha+(\beta+\gamma) \end{aligned}$$

   So we have $(\alpha+\beta)+\gamma = \alpha+(\beta+\gamma)$

2. **(commutative)** If $\alpha, \beta \in \mathbb{R}$ then we have

$$\begin{aligned} z \in \alpha+\beta &\Leftrightarrow z = r+s \wedge r \in \alpha \wedge s \in \beta \\ &\underset{7.7}{\Leftrightarrow} z = s+r \wedge r \in \alpha \wedge s \in \beta \\ &\Leftrightarrow z \in \beta+\alpha \end{aligned}$$

   giving $\alpha+\beta = \beta+\alpha$

3. **(neutral element)** Let $\alpha \in \mathbb{R}$. If $z \in \alpha + \alpha_0$ then $z = r+s$ where $r \in \alpha$ and $s < 0 \Rightarrow r+s < r \Rightarrow r+s \leqslant r \underset{8.3}{\Rightarrow} (r+s) \in \alpha \Rightarrow z \in \alpha \Rightarrow (\alpha+\alpha_0) \subseteq \alpha$. If $z \in \alpha$ then as α has no maximum (8.1 4), there exists a $z' \in \alpha$ with $z < z'$, then we have $z - z' < 0 \Rightarrow z - z' \in \alpha_0 \Rightarrow z = z' + (z - z') \in \alpha + \alpha_0 \Rightarrow \alpha \subseteq (\alpha+\alpha_0)$. This gives then $\alpha = \alpha+\alpha_0$ and by commutativity we have also $\alpha = \alpha_0+\alpha$.

4. **(inverse)** Let $\alpha \in \mathbb{R}$ take then $-\alpha$ and consider the following possible cases for $\min(\mathbb{Q}\backslash\alpha)$

   a. **(min $(\mathbb{Q}\backslash\alpha)$ does not exists)** If $x \in \alpha_0$ then $x < 0$ and thus $0 < -x$ using 8.11 there exists a $r \in \alpha$ such that $r - x = r + (-x) \notin \alpha$, as $\min(\mathbb{Q}\backslash\alpha)$ does not exists, we have $r - x \neq \min(\mathbb{Q}\backslash\alpha)$. Thus by definition of $-\alpha$ we have $-(r-x) \in -\alpha \Rightarrow x - r \in -\alpha \Rightarrow x = r + (x - r) \in \alpha + (-\alpha) \Rightarrow \alpha_0 \subseteq \alpha + (-\alpha)$. Also if $x \in \alpha + (-\alpha)$ then $x = r+s$ where $r \in \alpha$ and $s \in -\alpha \Rightarrow -s \notin \alpha$, using 8.1 3 we have then that $r < -s \Rightarrow r + s < 0 \Rightarrow x = r + s \in \alpha_0 \Rightarrow \alpha + (-\alpha) \subseteq \alpha_0$. Combining our two results gives $\alpha_0 = \alpha + (-\alpha)$.

   b. **(min $(\mathbb{Q}\backslash\alpha)$ exists)** If $r = \min(\mathbb{Q}\backslash\alpha)$ exist then by 8.4 we have that $\alpha = \alpha_r$, by 8.9 we have $-\alpha = -\alpha_r = \alpha_{-r}$. So we have $z \in \alpha + (-\alpha) = \alpha_r + \alpha_{-r}$ then there exists a $s < r$ and a $t < -r$ such that $z = s+t$, this gives $s+t < r+t$ and $t+r < r+(-r) = 0 \Rightarrow s+t < 0 \Rightarrow z < 0 \Rightarrow z \in \alpha_0 \Rightarrow \alpha+(-\alpha) \subseteq \alpha_0$. If $z \in \alpha_0 \Rightarrow z < 0 \Rightarrow 0 < -z$ and as $0 < \frac{1}{2}$ we have by 7.28 that $0 < (-z) \cdot \frac{1}{2} = -\left(z \cdot \frac{1}{2}\right) \Rightarrow \left(z \cdot \frac{1}{2}\right) < 0$. From this it follows that $r + \left(z \cdot \frac{1}{2}\right) < r \Rightarrow \left(r + \left(z \cdot \frac{1}{2}\right)\right) \in \alpha_r$ and $-r + \left(z \cdot \frac{1}{2}\right) < -r \Rightarrow \left(-r + \left(z \cdot \frac{1}{2}\right)\right) \in \alpha_{-r}$. Finally $z = \left(\frac{1}{2} \cdot z\right) + \left(\frac{1}{2} \cdot z\right) = \left(r + \left(z \cdot \frac{1}{2}\right)\right) + \left(-r + \left(z \cdot \frac{1}{2}\right)\right) \in \alpha_r + \alpha_{-r} = \alpha + (-\alpha) \Rightarrow \alpha_0 \subseteq \alpha + (-\alpha)$. Combining our two results gives $\alpha_0 = \alpha + (-\alpha)$

Using commutativity we have the also $\alpha_0 = (-\alpha) + \alpha$ $\square$

Next we define the set of positive and negative numbers

8.2.2 Multiplication in $\mathbb{R}$

Definition 8.14. *The set of positive real numbers $\mathbb{R}_+$ is defined by $\mathbb{R}_+ = \{\alpha \in \mathbb{R} | 0 < \alpha\} \subseteq \mathbb{R}$, the set of negative real numbers $\mathbb{R}_-$ is defined by $\mathbb{R}_- = \{\alpha | -\alpha \in \mathbb{R}_+\}$*

Theorem 8.15. $\mathbb{R} = \mathbb{R}_+ \bigcup \mathbb{R}_- \bigcup \{0\}$ *and* $\mathbb{R}_+ \bigcap \{0\} = \emptyset = \mathbb{R}_- \bigcap \{0\}$ *and* $\mathbb{R}_+ \bigcap \mathbb{R}_- = 0$*. In other words if $\alpha \in \mathbb{R}$ then we have either only one of $\alpha \in \mathbb{R}_+$ or $\alpha \in \mathbb{R}_-$ or $\alpha = 0$.*

Proof. First we prove that $\mathbb{R} = \mathbb{R}_+ \bigcup \mathbb{R}_- \bigcup \{0\}$, as $\mathbb{R}_+, \mathbb{R}_-, \{0\} \subseteq \mathbb{R}$ we have $\mathbb{R}_+ \bigcup \mathbb{R}_- \bigcup \{0\} \subseteq \mathbb{R}$. Now if $\alpha \in \mathbb{R}$ then we have the following possible cases

1. $(\mathbf{0} \in \boldsymbol{\alpha})$ $\Leftrightarrow \alpha \in \mathbb{R}_+$
2. $(\mathbf{0} \notin \boldsymbol{\alpha})$ here we have again two possible cases
   a. $(\mathbf{0} = \mathbf{min}\,(\mathbb{Q} \backslash \boldsymbol{\alpha}))$ $\underset{8.4}{\Leftrightarrow} \alpha = \alpha_0 = 0 \Leftrightarrow \alpha \in \{0\}$
   b. $(\mathbf{0} \neq \mathbf{min}\,(\mathbb{Q} \backslash \boldsymbol{\alpha}))$ $\underset{0 \in \alpha \,\text{and definition of}\, -\alpha \,(\text{see}\, 8.8)}{\Leftrightarrow} 0 = -0 \in -\alpha \Leftrightarrow -\alpha \in \mathbb{R}_+ \Leftrightarrow \alpha \in \mathbb{R}_-$

from (1),(2.a),(2.b) we conclude that $\mathbb{R} \subseteq \mathbb{R}_+ \bigcup \mathbb{R}_- \bigcup \{0\}$ and thus we have $\mathbb{R} = \mathbb{R}_+ \bigcup \mathbb{R}_- \bigcup \{0\}$.

Now if $0 \in \mathbb{R}_+$ then $0 = \alpha_0 \in \mathbb{R}_+ \Rightarrow 0 \in \alpha_0 \Rightarrow 0 < 0$ a contradiction, so we have $\{0\} \bigcap \mathbb{R}_+ = \emptyset$.

If $0 \in \mathbb{R}_-$ then $-0 \in \mathbb{R}_+ \Rightarrow_{-0 \underset{\text{neutral element}}{=} -0+0 \underset{\text{inverse element}}{=} 0} 0 \in \mathbb{R}_+$ which we have already prove to be impossible.

Finally if $\alpha \in \mathbb{R}_+ \bigcap \mathbb{R}_-$ then $\alpha \in \mathbb{R}_+ \wedge \alpha \in \mathbb{R}_- \Rightarrow \alpha \in \mathbb{R}_+ \wedge -\alpha \in \mathbb{R}_+ \Rightarrow 0 \in \alpha \wedge 0 \in -\alpha \Rightarrow 0 \in \alpha \wedge -0 \in -\alpha$ which is a contradiction by the definition of the negative (see 8.8). $\square$

Let's now proceed to definition of multiplication in $\mathbb{R}$, first we define multiplication on the set of positive reals

Definition 8.16. $\forall \alpha, \beta \in \mathbb{R}_+$ *then* $\alpha \odot \beta = \{r \in \mathbb{Q} | r \leqslant 0\} \bigcup \{s \cdot t | s \in \alpha \wedge t \in \beta, s > 0 \wedge t > 0\}$

We have then the following theorem

Theorem 8.17. $\forall \alpha, \beta \in \mathbb{R}_+$ *we have that* $\alpha \odot \beta \in \mathbb{R}_+$

Proof. Let's first prove that $\alpha \odot \beta \in \mathbb{R}$

1. $(\boldsymbol{\alpha} \odot \boldsymbol{\beta} \neq \boldsymbol{\emptyset})$ As $0 \in \{r \in \mathbb{Q} | r \leqslant 0\} \Rightarrow 0 \in \alpha \odot \beta \Rightarrow \alpha \odot \beta \neq \emptyset$

2. **$(\mathbb{Q}\backslash(\alpha\odot\beta)\neq\emptyset)$** As $\alpha,\beta\in\mathbb{R}_+$ we have $0\in\alpha$ and $0\in\beta$, from the fact that α,β do not have a maximum we must then have the existence of a $s_1\in\alpha$ and a $t_1\in\beta$ such that $0<s_1,0<t_1$. As $1=\frac{1}{1}\in\mathbb{Q}\backslash\{0\}$ there exists by 8.11 a $s_2\in\alpha$ and a $t_2\in\beta$ such that $s_2+1\notin\alpha$ and $t_2+1\in\beta$. Take $s=\max(s_1,s_2),t=\max(t_1,t_2)$ then $s\in\alpha,t\in\beta$ [if $s\notin\alpha\underset{8.1\ 3}{\Rightarrow}s_1<s,s_2<s\Rightarrow s\notin\{s_1,s_2\}$ a contradiction, if $t\notin\beta\underset{8.1\ 3}{\Rightarrow}t_1<t,t_2<t\Rightarrow t\notin\{t_1,t_2\}$ a contradiction] and $0<s,0<t$ and $s_2+1\leqslant s+1,t_2+1\leqslant t+1$ [as $s_2\leqslant\max(s_1,s_2)=s,t_2\leqslant\max(t_1,t_2)=t$]. We have then also $s+1\notin\alpha,t+1\notin\beta$ [if $s+1\in\alpha$ then by 8.1 3 and $s_2+1\notin\alpha$ we have the contradiction $s+1<s_2+1\leqslant s+1\Rightarrow s+1<s+1$, if $t+1\in\beta$ then by 8.1 3 and $t_2+1\notin\alpha$ we have the contradiction $t+1<t_2+1\leqslant t+1\Rightarrow t+1<t+1$]. We claim now that $s\cdot t+s+t+1\notin\alpha\odot\beta$ by contradiction, so let $s\cdot t+s+t\in\alpha\odot\beta$. First $0<s,0<t\underset{7.28}{\Rightarrow}0<s\cdot t$ and thus $0<s\cdot t+s\Rightarrow 0<s\cdot t+s+t\Rightarrow 0<s\cdot t+s+t+1\Rightarrow s\cdot t+s+t+1\notin\{r\in\mathbb{Q}|r\leqslant 0\}$, from this it follows that then we must have $s\cdot t+s+t+1\in\{s\cdot t|s\in\alpha\wedge t\in\beta\wedge 0<s\wedge 0<t\}$ so there exists a $s'\in\alpha,0<s'$ and a $t'\in\beta,0<t'$ such that $s\cdot t+s+t+1=s'\cdot t'$. Using 8.1 3 and $s+1\notin\alpha,t+1\notin\beta$ we have that $s'<s+1$, $t'<t+1\underset{0<s',0<t'<t+1}{\Rightarrow}s'\cdot t'<(s+1)\cdot t',(s+1)\cdot t=t'\cdot(s+1)<(s+1)\cdot(t+1)\Rightarrow s'\cdot t'<(s+1)\cdot(t+1)=s\cdot t+s+t+1$ and this gives then the contradiction $s\cdot t+s+t+1<s\cdot t+s+t+1$. So we conclude that $s\cdot t+s+t+1\notin\alpha\odot\beta\Rightarrow s\cdot t+s+t+1\in\mathbb{Q}\backslash(\alpha\odot\beta)\Rightarrow\mathbb{Q}\backslash(\alpha\odot\beta)\neq\emptyset$

3. If $r\in\alpha\odot\beta$ and $s\in\mathbb{Q}\backslash(\alpha\odot\beta)$then we have either

   a. **$(r\in\{s\in\mathbb{Q}|s\leqslant 0\})$** $\Rightarrow r\leqslant 0$ now for $s\in\mathbb{Q}\backslash(\alpha\odot\beta)$ we have $s\notin\{t\in\mathbb{Q}|t\leqslant 0\}\Rightarrow 0<s\Rightarrow r<s$

   b. **$(r\in\{s\in\mathbb{Q}|s\leqslant 0|\})$** $\Rightarrow r\in\{s\cdot t|s\in\alpha\wedge t\in\beta\wedge 0<s\wedge 0<t\}$ and $0\leqslant r$ so $r=s'\cdot t'$ where $s'\in\alpha\wedge t'\in\beta\wedge s'>0\wedge t'>0$ so we have by 7.28 that $0<s'\cdot t'=r$. Suppose now that $s\in\mathbb{Q}\backslash(\alpha\odot\beta)$ and that $s\leqslant r$, we will derive a contradiction from this. As $s\notin\alpha\odot\beta$ we have $s\neq r$ and $s\notin\{s\in\mathbb{Q}|s\leqslant 0\}\Rightarrow 0<s$ and thus $0<s<r$ using 7.10 we have that s^{-1} exists and by 7.34 $0<s^{-1}$. Take now $t=s^{-1}\cdot r\Rightarrow s\cdot t=r$ then we have $0<t$, using 7.28 and $s<r$ this gives $s^{-1}s<s^{-1}\cdot r=t\Rightarrow 1<t$. As $r=s\cdot t=s'\cdot t'$ and $0<1<t$ we have $s=t^{-1}\cdot(s'\cdot t')$. From $1<t$ we have by 7.35 that $t^{-1}<1$ and by 7.35 and $0<s'$ we get $t^{-1}\cdot s'<s'$ and thus $t^{-1}\cdot s'\in\alpha$ [if $t^{-1}\cdot s'\in\alpha\underset{8.1,3}{\Rightarrow}s'\leqslant t^{-1}\cdot s'<s'\Rightarrow s'<s'$ a contradiction]. From $t^{-1}\cdot s'\in\alpha$ and $t'\in\beta$ we have $s=(t^{-1}\cdot s')\cdot t'\in\alpha\odot\beta$ contradiction $s\in\mathbb{Q}\backslash(\alpha\odot\beta)$ so we must have $r<s$.

4. We prove now by contradiction that $\max(\alpha\odot\beta)$ does not exists. So assume that $m=\max(\alpha\odot\beta)$ exists then as $m\in\alpha\odot\beta$ we have to consider the following two cases

   a. **$(m\in\{s\in\mathbb{Q}|s\leqslant 0\})$** $\Rightarrow m\leqslant 0$, now $\alpha,\beta\in\mathbb{R}_+\Rightarrow 0\in\alpha,0\in\beta$ and as α,β don't have a maximum there exists a $s'\in\alpha,0<s'$ and a $t'\in\beta$, $0<\beta\Rightarrow 0<s'\cdot t'\in\alpha\odot\beta$ and thus $m<s'\cdot t'$ contradicting the fact that $m=\max(\alpha\odot\beta)$, so we reach a contradiction.

b. $\boldsymbol{(m\in\{s\in\mathbb{Q}|s\leqslant 0\})}$ $\Rightarrow 0<m$ and $m\in\{s\cdot t|s\in\alpha\wedge t\in\beta\wedge 0<s\wedge 0<t\}\Rightarrow\exists s\in\alpha, 0<s, \exists t\in\beta, 0<t$ with $m=s\cdot t$. As $\max(\alpha)$ and $\max(\beta)$ does not exists we have that $\exists s'\in\alpha\vdash 0<s<s', \exists t'\in\beta\vdash 0<t<t'\Rightarrow m=s\cdot t<s'\cdot t'\in\alpha\odot\beta$ contradicting the maximality of m. $\square$

Using the above definition we can define multiplication on $\mathbb{R}$, first we prove the following lemma.

Lemma 8.18. $\mathbb{R}\times\mathbb{R}$ *is the disjoint union of* $\mathbb{R}_+\times\mathbb{R}_+, \mathbb{R}_-\times\mathbb{R}_-, \mathbb{R}_+\times\mathbb{R}_-, \mathbb{R}_-\times\mathbb{R}_+, (\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\})$

Proof. First as $\mathbb{R}_+, \mathbb{R}_-, \{0\}\subseteq\mathbb{R}$ and $(\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\})\subseteq\mathbb{R}\times\mathbb{R}$ we have $(\mathbb{R}_+\times\mathbb{R}_+)\bigcup(\mathbb{R}_-\times\mathbb{R}_-)\bigcup(\mathbb{R}_+\times\mathbb{R}_-)\bigcup(\mathbb{R}_-\times\mathbb{R}_+)\bigcup((\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\}))\subseteq\mathbb{R}\times\mathbb{R}$. Second if $(x,y)\in\mathbb{R}\times\mathbb{R}$ then using 8.15 we have $(x\in\mathbb{R}_+\vee x\in\mathbb{R}_-\vee x=0)\wedge(y\in\mathbb{R}_+\vee y\in\mathbb{R}_-\vee y=0)\Rightarrow(x\in\mathbb{R}_+\wedge y\in\mathbb{R}_+)\vee(x\in\mathbb{R}_+\wedge y\in\mathbb{R}_-)\vee(x\in\mathbb{R}_+\vee y=0)\vee(x\in\mathbb{R}_-\wedge y\in\mathbb{R}_+)\vee(x\in\mathbb{R}_-\wedge y\in\mathbb{R}_-)\vee(x\in\mathbb{R}_-\vee y=0)\vee(x\in 0\wedge y\in\mathbb{R}_+)\vee(x\in 0\wedge y\in\mathbb{R}_-)\vee(x\in 0\vee y=0)\Rightarrow(x,y)\in(\mathbb{R}_+\times\mathbb{R}_+)\bigcup(\mathbb{R}_-\times\mathbb{R}_-)\bigcup(\mathbb{R}_+\times\mathbb{R}_-)\bigcup(\mathbb{R}_-\times\mathbb{R}_+)\bigcup((\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\}))\Rightarrow\mathbb{R}\times\mathbb{R}\subseteq(\mathbb{R}_+\times\mathbb{R}_+)\bigcup(\mathbb{R}_-\times\mathbb{R}_-)\bigcup(\mathbb{R}_+\times\mathbb{R}_-)\bigcup(\mathbb{R}_-\times\mathbb{R}_+)\bigcup((\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\}))$. This proves that $(\mathbb{R}_+\times\mathbb{R}_+)\bigcup(\mathbb{R}_-\times\mathbb{R}_-)\bigcup(\mathbb{R}_+\times\mathbb{R}_-)\bigcup(\mathbb{R}_-\times\mathbb{R}_+)\bigcup((\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\}))=\mathbb{R}\times\mathbb{R}$. Next to prove that the union is a disjoint union. From the fact that $\mathbb{R}$ is the disjoint union of $\mathbb{R}_+, \mathbb{R}_-$ and $\{0\}$ we have the following

1. If $(x,y)\in(\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\})$ then $x=0\vee y=0$ so $(x,y)\notin\mathbb{R}_+\times\mathbb{R}_+, \mathbb{R}_+\times\mathbb{R}_-, \mathbb{R}_-\times\mathbb{R}_+, \mathbb{R}_-\times\mathbb{R}_-$ and thus $((\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\}))\bigcap(\mathbb{R}_+\times\mathbb{R}_+)=\emptyset$, $((\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\}))\bigcap(\mathbb{R}_+\times\mathbb{R}_-)=\emptyset, ((\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\}))\bigcap(\mathbb{R}_-\times\mathbb{R}_+)=\emptyset, ((\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\}))\bigcap(\mathbb{R}_-\times\mathbb{R}_-)=\emptyset$

2. If $(x,y)\in\mathbb{R}_+\times\mathbb{R}_+$ then $x\neq 0, x\notin\mathbb{R}_-, y\neq 0, y\notin\mathbb{R}_-$ so $(x,y)\notin(\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\}), \mathbb{R}_+\times\mathbb{R}_-, \mathbb{R}_-\times\mathbb{R}_+, \mathbb{R}_-\times\mathbb{R}_-$ and thus $(\mathbb{R}_+\times\mathbb{R}_+)\bigcap(\mathbb{R}_+\times\mathbb{R}_-)=\emptyset, (\mathbb{R}_+\times\mathbb{R}_+)\bigcap(\mathbb{R}_-\times\mathbb{R}_+)=\emptyset, (\mathbb{R}_+\times\mathbb{R}_+)\bigcap(\mathbb{R}_-\times\mathbb{R}_-)=\emptyset, (\mathbb{R}_+\times\mathbb{R}_+)\bigcap((\{0\}\times\mathbb{R})\bigcup(\{\mathbb{R}\times\{0\}\}))=\emptyset$

3. If $(x,y)\in\mathbb{R}_+\times\mathbb{R}_-$ then $x\neq 0, x\notin\mathbb{R}_-, y\neq 0, y\notin\mathbb{R}_+$ so $(x,y)\notin(\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\}), \mathbb{R}_+\times\mathbb{R}_+, \mathbb{R}_-\times\mathbb{R}_+, \mathbb{R}_-\times\mathbb{R}_-$ and thus $(\mathbb{R}_+\times\mathbb{R}_-)\bigcap(\mathbb{R}_+\times\mathbb{R}_+)=\emptyset, (\mathbb{R}_+\times\mathbb{R}_-)\bigcap(\mathbb{R}_-\times\mathbb{R}_+)=\emptyset, (\mathbb{R}_+\times\mathbb{R}_-)\bigcap(\mathbb{R}_-\times\mathbb{R}_-)=\emptyset, (\mathbb{R}_+\times\mathbb{R}_+)\bigcap((\{0\}\times\mathbb{R})\bigcup(\{\mathbb{R}\times\{0\}\}))=\emptyset$

4. If $(x,y)\in\mathbb{R}_-\times\mathbb{R}_+$ then $x\neq 0, x\notin\mathbb{R}_+, y\neq 0, y\notin\mathbb{R}_-$ so $(x,y)\notin(\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\}), \mathbb{R}_+\times\mathbb{R}_-, \mathbb{R}_+\times\mathbb{R}_+, \mathbb{R}_-\times\mathbb{R}_-$ and thus $(\mathbb{R}_-\times\mathbb{R}_+)\bigcap(\mathbb{R}_+\times\mathbb{R}_-)=\emptyset, (\mathbb{R}_-\times\mathbb{R}_+)\bigcap(\mathbb{R}_+\times\mathbb{R}_-)=\emptyset, (\mathbb{R}_+\times\mathbb{R}_-)\bigcap(\mathbb{R}_-\times\mathbb{R}_-)=\emptyset, (\mathbb{R}_+\times\mathbb{R}_-)\bigcap((\{0\}\times\mathbb{R})\bigcup(\{\mathbb{R}\times\{0\}\}))=\emptyset$

5. If $(x,y)\in\mathbb{R}_-\times\mathbb{R}_-$ then $x\neq 0, x\notin\mathbb{R}_+, y\neq 0, y\notin\mathbb{R}_+$ so $(x,y)\notin(\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\}), \mathbb{R}_+\times\mathbb{R}_+, \mathbb{R}_+\times\mathbb{R}_-, \mathbb{R}_-\times\mathbb{R}_+$ and thus $(\mathbb{R}_-\times\mathbb{R}_-)\bigcap(\mathbb{R}_+\times\mathbb{R}_+)=\emptyset, (\mathbb{R}_-\times\mathbb{R}_-)\bigcap(\mathbb{R}_+\times\mathbb{R}_-)=\emptyset, (\mathbb{R}_-\times\mathbb{R}_-)\bigcap(\mathbb{R}_-\times\mathbb{R}_+)=\emptyset, (\mathbb{R}_-\times\mathbb{R}_-)\bigcap((\{0\}\times\mathbb{R})\bigcup(\{\mathbb{R}\times\{0\}\}))=\emptyset$ $\square$

Using the previous lemma and 2.31 we can define $\cdot:\mathbb{R}\times\mathbb{R}\rightarrow\mathbb{R}$

Definition 8.19. *We define* $\cdot : \mathbb{R}\times\mathbb{R}\to\mathbb{R}$ $(\alpha,\beta)\to\alpha\cdot\beta$ *as follows*

$$\alpha\cdot\beta=\begin{cases}\alpha\odot\beta\,\mathrm{if}\,(\alpha,\beta)\in\mathbb{R}_+\times\mathbb{R}_+\\ -((-\alpha)\odot\beta)\,\mathrm{if}\,(\alpha,\beta)\in\mathbb{R}_-\times\mathbb{R}_+\\ -(\alpha\times(-\beta))\,\mathrm{if}\,(\alpha,\beta)\in\mathbb{R}_+\times\mathbb{R}_-\\ (-\alpha)\odot(-\beta)\,\mathrm{if}\,(\alpha,\beta)\in\mathbb{R}_+\times\mathbb{R}_+\\ 0\,\mathrm{if}\,(\alpha,\beta)\in(\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\})\end{cases}$$

If we want to prove something about multiplication we have to consider each time 5 different cases, luckily the following lemma will allow us to bring down the different cases.

Lemma 8.20. $\forall(\alpha,\beta)\in\mathbb{R}\times\mathbb{R}$ *we have* $-(\alpha\cdot\beta)=(-\alpha)\cdot\beta=\alpha\cdot(-\beta)$

Proof. We have to consider the following 5 cases

1. $\boldsymbol{((\alpha,\beta)\in\mathbb{R}_+\times\mathbb{R}_+)}$ We have then

   a. $-\alpha\in\mathbb{R}_-$ and (see 3.6) $\alpha=-(-\alpha)$ then we have

   $$\begin{aligned}-(\alpha\cdot\beta)&=-(\alpha\odot\beta)\\&=-(-(-\alpha)\odot\beta)\\&=-(-(-\alpha\cdot\beta))\\&=-(-\alpha\cdot\beta)\end{aligned}$$

   b. $-\beta\in\mathbb{R}_-$ and (see 3.6) $\beta=-(-\beta)$ then we have

   $$\begin{aligned}-(\alpha\cdot\beta)&=-(\alpha\odot\beta)\\&=-(\alpha\odot(-(-\beta)))\\&=-(-(\alpha\cdot(-\beta)))\\&=(\alpha\cdot(-\beta))\end{aligned}$$

2. $\boldsymbol{((\alpha,\beta)\in\mathbb{R}_+\times\mathbb{R}_-)}$ We have then

   a. $-\alpha\in\mathbb{R}_-$ and $\alpha=-(-\alpha)$ then we have

   $$\begin{aligned}-(\alpha\cdot\beta)&=-(-(\alpha\odot(-\beta)))\\&=(-(-\alpha)\odot(-\beta))\\&=((-\alpha)\cdot\beta)\end{aligned}$$

   b. $-\beta\in\mathbb{R}_+$ and $\beta=-(-\beta)$ then we have

   $$\begin{aligned}-(\alpha\cdot\beta)&=-(-(\alpha\odot(-\beta)))\\&=(\alpha\odot(-\beta))\\&=(\alpha\cdot(-\beta))\end{aligned}$$

3. $\boldsymbol{((\alpha,\beta)\in\mathbb{R}_-\times\mathbb{R}_+)}$ We have then

   a. $-\alpha\in\mathbb{R}_+$ and $\alpha=-(-\alpha)$ then we have

   $$\begin{aligned}-(\alpha\cdot\beta)&=-(-(-\alpha\odot\beta))\\&=(-\alpha\cdot\beta)\end{aligned}$$

b. $-\beta \in \mathbb{R}_-$ and $\beta = -(-\beta)$ then we have

$$\begin{array}{rcl} -(\alpha \cdot \beta) & = & -(-(-\alpha \odot \beta)) \\ & = & (-\alpha \odot \beta) \\ & = & (-\alpha \odot (-(-\beta))) \\ & = & (\alpha \cdot (-\beta)) \end{array}$$

4. $\boldsymbol{((\alpha, \beta) \in \mathbb{R}_- \times \mathbb{R}_-)}$ We have then

a. $-\alpha \in \mathbb{R}_+$ and $\alpha = -(-\alpha)$ then we have

$$\begin{array}{rcl} -(\alpha \cdot \beta) & = & -(-\alpha \odot (-\beta)) \\ & = & -(-(-\alpha \cdot \beta)) \\ & = & (-\alpha \cdot \beta) \end{array}$$

b. $-\beta \in \mathbb{R}_+$ and $\beta = -(-\beta)$ then we have

$$\begin{array}{rcl} -(\alpha \cdot \beta) & = & -(-\alpha \odot -\beta) \\ & = & -(-(\alpha \cdot (-\beta))) \\ & = & (\alpha \cdot (-\beta)) \end{array}$$

5. $\boldsymbol{((\alpha, \beta) \in (\{0\} \times \mathbb{R}) \bigcup (\mathbb{R} \times \{0\}))}$ then we have the following (non exclusive cases)

a. $\alpha = 0$ We have then

$$\begin{array}{rcl} -(\alpha \cdot \beta) & = & -(0 \cdot \beta) \\ & = & 0 \\ & = & (0 \cdot \beta) \\ & = & (-0 \cdot \beta) \\ & = & (-\alpha \cdot \beta) \\ -(\alpha \cdot \beta) & = & -(0 \cdot \beta) \\ & = & 0 \\ & = & (0 \cdot (-\beta)) \\ & = & (\alpha \cdot (-\beta)) \end{array}$$

b. $\beta = 0$ We have then

$$\begin{array}{rcl} -(\alpha \cdot \beta) & = & -(\alpha \cdot 0) \\ & = & 0 \\ & = & (-\alpha, 0) \\ & = & (-\alpha, \beta) \\ -(\alpha \cdot \beta) & = & -(\alpha \cdot 0) \\ & = & 0 \\ & = & (\alpha \cdot (-0)) \\ & = & (\alpha \cdot (-\beta)) \end{array}$$

$\square$

Theorem 8.21. *$\forall \alpha \in \mathbb{R}_+$ we have that* $\operatorname{rep}(\alpha) = \{r \in \mathbb{Q} | r \leqslant 0\} \bigcup \{s^{-1} | s \in \mathbb{Q} \backslash \alpha \wedge s > 0 \wedge s \neq \min(\mathbb{Q} \backslash \alpha)\}$ *is a cut. This is called the reciprocal cut of α.*

Proof. We have

1. **($\operatorname{rep}(\alpha) \neq \emptyset$)** As $0 \in \{r \in \mathbb{Q} | r \leqslant 0\}$ we have $0 \in \operatorname{rep}(\alpha) \Rightarrow \operatorname{rep}(\alpha) \neq \emptyset$
2. **($\operatorname{rep}(\alpha) \neq \mathbb{Q}$)** As $0 \in \alpha$ then from the fact that there does not exists a maximum for α there exists a $s \in \alpha$ with $s > 0$. Using 7.34 we have $0 < s^{-1} \Rightarrow s^{-1} \notin \{r \in \mathbb{Q} | r \leqslant 0\}$ and as $s \in \alpha$ we have $s^{-1} \notin \{s^{-1} | s \in \mathbb{Q} \backslash \alpha \wedge s > 0 \wedge s \neq \min(\mathbb{Q} \backslash \alpha)\}$ so $s^{-1} \in \mathbb{Q} \backslash \operatorname{rep}(\alpha)$
3. **($\forall r \in \operatorname{rep}(\alpha) \wedge \forall s \in \mathbb{Q} \backslash \operatorname{rep}(\alpha)$ we have $r < s$)** Let $r \in \operatorname{rep}(\alpha)$ and $s \in \mathbb{Q} \backslash \operatorname{rep}(\alpha) \Rightarrow s \notin \operatorname{rep}(\alpha)$. We have for r the following exclusive cases
   a. **($r \leqslant 0$)** then as $s \notin \operatorname{rep}(\alpha)$ we must have $s \in \{r \in \mathbb{Q} | r \leqslant 0\} \Rightarrow 0 < s \Rightarrow r < s$
   b. **($0 < r$)** then $0 < r^{-1}$, $r^{-1} \in \mathbb{Q} \backslash \alpha \Rightarrow r^{-1} \notin \alpha$ and $r^{-1} \neq \min(\mathbb{Q} \backslash \alpha)$. As $s \in \mathbb{Q} \backslash \operatorname{rep}(\alpha)$ we have $0 < s \Rightarrow 0 < s^{-1}$. Consider now the following possible cases for s^{-1} then
      i. **($s^{-1} = \min(\mathbb{Q} \backslash \alpha)$)** then as $r^{-1} \notin \alpha \Rightarrow r^{-1} \notin \mathbb{Q} \backslash \alpha \underset{s^{-1} = \min(\mathbb{Q} \backslash \alpha)}{\Rightarrow} s^{-1} \leqslant r^{-1} \underset{r^{-1} \neq \min(\mathbb{Q} \backslash \alpha)}{\Rightarrow} s^{-1} < r^{-1} \underset{7.35}{\Rightarrow} r < s$
      ii. **($s^{-1} \neq \min(\mathbb{Q} \backslash \alpha)$)** As $s \notin \operatorname{rep}(\alpha)$ we have $0 < s$ [$s \notin \{r \in \mathbb{Q} | r \leqslant 0\}$] and thus by 7.35 we have $0 < s^{-1}$. We must then have $s^{-1} \in \alpha$ [otherwise $s^{-1} \in \{s^{-1} | s \in \mathbb{Q} \backslash \alpha \wedge s > 0 \wedge s \neq \min(\mathbb{Q} \backslash \alpha)\} \subseteq \operatorname{rep}(\alpha)$] and using 8.1 , 3 we have then $s^{-1} < r^{-1} \underset{7.35}{\Rightarrow} r < s$
4. **($\operatorname{rep}(\alpha)$ does not have a greatest element)** We prove this by contradiction. So assume that $m = \max(\operatorname{rep}(\alpha)) \in \operatorname{rep}(\alpha)$ exists, then we we derive a contradiction from this. We can now have for $m \in \operatorname{rep}(\alpha)$ the following possibilities
   a. **($m \leqslant 0$)** Using 8.1 , 2 there exists a $s \in \mathbb{Q} \backslash \alpha \Rightarrow s \notin \alpha$, as $\alpha \in \mathbb{R}_+ \Rightarrow 0 \in \alpha \underset{8.1,\ 3}{\Rightarrow} 0 < s \Rightarrow 0 < s < s+1$. From $0 < s < s+1$ we have $s+1 \notin \alpha$ [for otherwise a s $s \notin \alpha$ we have by 8.1 , 3 that $s+1 < s$ contradicting $s < s+1$], as $s < s+1$ we have $s+1 \neq \min(\mathbb{Q} \backslash \alpha)$ and from $0 < s+1$ we have $0 < (s+1)^{-1}$ and thus we have $(s+1)^{-1} \in \operatorname{rep}(\alpha)$. Then $0 \leqslant m < (s+1)^{-1} \leqslant m \Rightarrow m < m$ and we reach a contradiction.
   b. **($0 < m$)** then as $m \in \operatorname{rep}(\alpha)$ we must have $0 < m^{-1}, m^{-1} \notin \alpha$ and $m^{-1} \neq \min(\mathbb{Q} \backslash \alpha)$. From the last inequality we have the existence of a $t \in \mathbb{Q} \backslash \alpha$ such that $t < m^{-1}$. Using the fact that $\mathbb{Q}$ is dense (see 7.36) there exists a $s \in \mathbb{Q}$ such that $t < s < m^{-1}$ then we must have $s \notin \alpha$ [for otherwise by 8.1 , 3 I would have $s < t$ contradicting $t < s$]. As $\alpha \in \mathbb{R}_+ \Rightarrow 0 \in \alpha \Rightarrow 0 < t < s < m^{-1}$ then as $t \notin \alpha$ we must have $s \notin \alpha$ [otherwise we must have $s < t$ contradicting $t < s$]. Since $t < s$ and t, $s \in \mathbb{Q} \backslash \alpha$ we have $s \neq \min(\mathbb{Q} \backslash \alpha)$. We must then from $0 < s \Rightarrow 0 < s^{-1}$, $s \neq \min(\mathbb{Q} \backslash \alpha)$, $s \in \mathbb{Q} \backslash \alpha$ conclude that $s^{-1} \in \operatorname{rep}(\alpha)$. From $0 < s < m^{-1}$ we have $m = (m^{-1})^{-1} < s^{-1} \Rightarrow m < s^{-1}$ contradicting the maximality of m. □

Definition 8.22. $\mathbb{R}_0 = \mathbb{R}\backslash\{0\}$. *Note that by 8.15 we have* $\mathbb{R}_0 = \mathbb{R}_+ \bigcup \mathbb{R}_-$

We are now ready to define the multiplicative inverse of non zero real numbers.

Definition 8.23. *If* $r \in \mathbb{R}_0$ *the we define* r^{-1} *as follows* $r^{-1} = \begin{cases} \text{rep}(r) \text{ if } r \in \mathbb{R}_+ \\ -\text{rep}(-r) \text{ if } r \in \mathbb{R}_+ \end{cases}$

Actual the reciprocal cut of a cut takes a rather simple form in the case of positive rational cuts.

Lemma 8.24. *If* $r \in \mathbb{Q}$ *with* $\alpha_r \in \mathbb{R}_0$ *then* $r^{-1} = \alpha_{r^{-1}}$ *(note that there are two different inverses here one in* $\mathbb{Q}$ *and one in* $\mathbb{R}_0$*)*

Proof. First we take the case of $a_r \in \mathbb{R}_+$ then we have as $0 \in \alpha_r \Rightarrow 0 < r \Rightarrow 0 < r^{-1}$. For $x \in (\alpha_r)^{-1}$ the following possible cases

1. $(\boldsymbol{x} \leqslant \mathbf{0})$ then $x \leqslant 0 < r^{-1} \Rightarrow x \in \alpha_{r^{-1}}$
2. $(\mathbf{0} < \boldsymbol{x})$ then we must have $0 < x^{-1}, x^{-1} \in \mathbb{Q}\backslash\alpha_r, x^{-1} \neq \min(\mathbb{Q}\backslash\alpha_r) \underset{8.4}{=} r \Rightarrow r < x^{-1} \underset{7.35}{\Rightarrow} (x^{-1})^{-1} < r^{-1} \Rightarrow x < r^{-1} \Rightarrow x \in \alpha_{r^{-1}}$

so we have $(\alpha_r)^{-1} \subseteq \alpha_{r^{-1}}$

If $x \in \alpha_{r^{-1}}$ then we have the following possible cases for x

1. $(\boldsymbol{x} \leqslant \mathbf{0})$ Then $x \in (\alpha_r)^{-1}$ trivially.
2. $(\mathbf{0} < \boldsymbol{x})$ Then as $x \in \alpha_{r^{-1}} \Rightarrow x < r^{-1} \underset{7.35}{\Rightarrow} r < x^{-1} \Rightarrow r = \min(\mathbb{Q}\backslash\alpha_r) \neq x^{-1}$, $0 < x^{-1}$ and $x^{-1} \in \alpha_r \Rightarrow x^{-1} \in \mathbb{Q}\backslash\alpha_r \Rightarrow x \in (\alpha_r)^{-1}$

so we have $\alpha_{r^{-1}} \in (\alpha_r)^{-1}$. And we conclude thus that $\alpha_{r^{-1}} = (\alpha_r)^{-1}$.

Second if $\alpha_r \in \mathbb{R}_-$ then $(\alpha_r)^{-1} = -(\text{rep}(-\alpha)) \underset{-\alpha \in \mathbb{R}_+ \text{and previous}}{=} -(\alpha_{(-r)^{-1}}) \underset{7.35}{=} -(\alpha_{-(r^{-1})}) \underset{8.9}{=} \alpha_{(-(-(r^{-1})))} = \alpha_{r^{-1}}$ □

Theorem 8.25. ($\mathbb{R}$ forms *a* field) $\langle \mathbb{R}, +, \cdot \rangle$ *forms a field. The neutral element (identity) from multiplication is* α_1*, to simplify our notations in the future we note* α_1 *as* 1 *(again* 1 *means different things in different fields). The multiplicative inverse of* α *is* α^{-1}

Proof. First we have already proved that $\langle \mathbb{R}, + \rangle$ is a abelian group (see 8.13) we prove then the following for multiplication. In the proof we make heavily use of the lemma 8.20.

1. **(Commutativity)** we have to consider the following five cases
   a. $((\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathbb{R}_+ \times \mathbb{R}_+)$ Here we have

$$\begin{aligned} \alpha \cdot \beta &= \alpha \odot \beta \\ &= \{r \in \mathbb{Q} | r \leqslant 0\} \bigcup \{s \cdot t | s \in \alpha, t \in \beta, s > 0, t > 0\} \\ &\underset{7.10}{=} \{r \in \mathbb{Q} | r \leqslant 0\} \bigcup \{t \cdot s | s \in \alpha, t \in \beta, s > 0, t > 0\} \\ &= \beta \odot \alpha \\ &= \beta \cdot \alpha \end{aligned}$$

   b. $((\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathbb{R}_+ \times \mathbb{R}_-)$ Here we have

$$\begin{aligned} \alpha \cdot \beta &= -(\alpha \cdot (-\beta)) \\ &\underset{1.a}{=} -(\beta \cdot (-\alpha)) \\ &= \beta \cdot \alpha \end{aligned}$$

c. $\boldsymbol{((\alpha,\beta)\in\mathbb{R}_-\times\mathbb{R}_+)}$ Here we have

$$\begin{array}{rcl}\alpha\cdot\beta & = & -(-\alpha\cdot\beta)\\ & \underset{1.a}{=} & -(\beta\cdot(-\alpha))\\ & = & \beta\cdot\alpha\end{array}$$

d. $\boldsymbol{((\alpha,\beta)\in\mathbb{R}_-\times\mathbb{R}_-)}$ Here we have

$$\begin{array}{rcl}\alpha\cdot\beta & = & -(-(-\alpha\cdot(-\beta)))\\ & = & ((-\alpha)\cdot(-\beta))\\ & \underset{1.a}{=} & ((-\beta)\cdot(-\alpha))\\ & = & \beta\cdot\alpha\end{array}$$

e. $\boldsymbol{((\alpha,\beta)\in((\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\})))}$ Here we have

$$\begin{array}{rcl}\alpha\cdot\beta & = & 0\\ & = & \beta\cdot\alpha\end{array}$$

2. **(Neutral element)**

First note that $0<1$ so $0\in\alpha_1$ we have the to consider the following cases

a. $\boldsymbol{(\alpha\in\mathbb{R}_+)}$ If $x\in\alpha\odot\alpha_1$ then we have either

i. $\boldsymbol{(x\leqslant 0)}$ As $\alpha\in\mathbb{R}_+\Rightarrow 0\in\alpha$ then using 8.3 we have $x\in\alpha$

ii. $\boldsymbol{(0<x)}$ So there exists a $s\in\alpha, t\in\alpha_1\Rightarrow t<1$ with $0<s,t$ and $x=s\cdot t$, from $t<1$ we have $x=s\cdot t<s\Rightarrow x\leqslant s$, as $s\in\alpha$ we have by 8.3 that $x\in\alpha$.

So we conclude that $\alpha\odot\alpha_1\subseteq\alpha$. If $x\in\alpha$ then we have the following possible cases

i. $\boldsymbol{(x\leqslant 0)}$ then $x\in\alpha\odot\alpha_1$

ii. $\boldsymbol{(0<x)}$ then as $\max(\alpha)$ does not exists there exists a $s\in\alpha$ such that $0<x<s$ then $0<x\cdot s^{-1}<1\Rightarrow 0<x\cdot s^{-1}\in\alpha_1$ and from $x=s\cdot(x\cdot s^{-1})$ we have $x\in\alpha\odot\alpha_1$

this gives $\alpha\subseteq\alpha\odot\alpha_1$. Or to summarize we have $\alpha=\alpha\odot\alpha_1$. Now we have

$$\begin{array}{rcl}\alpha & = & \alpha\odot\alpha_1\\ & = & \alpha\cdot\alpha_1\\ & \underset{\text{commutativity}}{=} & \alpha_1\cdot\alpha\end{array}$$

b. $\boldsymbol{(\alpha\in\mathbb{R}_-)}$ Here we have

$$\begin{array}{rcl}\alpha\cdot\alpha_1 & = & -((-\alpha)\cdot\alpha_1)\\ & \underset{2.a}{=} & -(-\alpha)\\ & = & \alpha\\ & \underset{\text{commutativity}}{=} & \alpha_1\cdot\alpha\end{array}$$

c. **($\boldsymbol{\alpha = 0}$)** Here we have

$$\begin{aligned} \alpha\cdot\alpha_1 &= 0 \\ &= \alpha_1\cdot\alpha \end{aligned}$$

3. **(Inverse for non zero element)** As $\alpha\in\mathbb{R}_0$ we have the following cases to consider

   a. **($\boldsymbol{\alpha\in\mathbb{R}_+}$)** Take then $x\in\alpha\odot\alpha^{-1}=\alpha\odot\text{rep}(\alpha)$ then we have either

      i. **($\boldsymbol{x\leqslant 0}$)** then as $0<1$ we have $x<1\Rightarrow x\in\alpha_1$

      ii. **($\boldsymbol{0<x}$)** then $\exists s\in\alpha\vdash 0<s$ and $\exists t\in\alpha^{-1}\vdash 0<t$ such that $s\cdot t=x$. Now for t we have the following cases

         A. **($\boldsymbol{t\leqslant 0}$)** then from $0<s$ we have $x=s\cdot t\leqslant 0<1\Rightarrow x<1\Rightarrow x\in\alpha_1$

         B. **($\boldsymbol{0<t}$)** then $t^{-1}\in\mathbb{Q}\backslash\alpha$, $0<t^{-1}$ and $t^{-1}\neq\min(\mathbb{Q}\backslash\alpha)$. As $s\in\alpha$ we have from $t^{-1}\in\mathbb{Q}\backslash\alpha$ and 8.1 , 3 that $s<t^{-1}\Rightarrow x=s\cdot t<1\Rightarrow x\in\alpha_1$

   Take now $x\in\alpha_1\Rightarrow x<1$ then we have either

      i. **($\boldsymbol{x\leqslant 0}$)** $\Rightarrow x\in\alpha\odot\text{rep}(\alpha)$

      ii. **($\boldsymbol{0<x}$)** then we have $x^{-1}\in\mathbb{Q}\backslash\alpha, 0<x^{-1}$ and $x^{-1}\neq\min(\mathbb{Q}\backslash\alpha)$. From $0<x^{-1}$ we have by 7.34 that $0<x\Rightarrow 0<x<1\Rightarrow 0<1-x$. Since $\alpha\in\mathbb{R}_+\Rightarrow 0\in\alpha$ and as $\max(\alpha)$ does not exists there exists a $s_1\in\alpha$ such that $0<s_1$. From $0<x<1$ we have $1^{-1}=1<x^{-1}$ and thus $0<1-x<x^{-1}(x-1)\Rightarrow 0<s_1\cdot(x-1)\cdot x^{-1}$. If we note $\varepsilon=s_1\cdot(x-1)\cdot x^{-1}$ and $0<\varepsilon$. Using 8.11 there exists a $s_2\in\alpha$ such that $\varepsilon+s_2\notin\alpha$. Then we can find a $s_3\in\alpha$ such that $\varepsilon+s_2\leqslant\varepsilon+s_3\in\mathbb{Q}\backslash\alpha, \varepsilon+s_3\neq\min(\mathbb{Q}\backslash\alpha)$ [[If $\varepsilon+s_2\neq\min(\mathbb{Q}\backslash\alpha)$ we take $s_3=s_2$, else if $\varepsilon+s_2=\min(\mathbb{Q}\backslash\alpha)$ then as $\max(\alpha)$ does not exists there exists a $s_3\in\alpha$ with $s_2<s_3\Rightarrow\varepsilon+s_2<\varepsilon+s_3$, now $\varepsilon+s_3\in\mathbb{Q}\backslash\alpha$ [otherwise $\varepsilon+s_3\in\alpha\underset{8.1}{\Rightarrow}\varepsilon+s_3<\varepsilon+s_2$ contradicting $\varepsilon+s_2<\varepsilon+s_3$], we have then $\varepsilon+s_3\neq\min(\mathbb{Q}\backslash\alpha)$]]. Take now $s=\max(s_1,s_3)$. Then as $0\in\alpha$ we have as $\varepsilon+s_2\in\mathbb{Q}\backslash\alpha$ by 8.1 that $0<\varepsilon+s_2\leqslant\varepsilon+s$, $\varepsilon+s\in\mathbb{Q}\backslash\alpha$ [if $\varepsilon+s\in\alpha\underset{\varepsilon+s_2\in\mathbb{Q}\backslash\alpha,8.1}{\Rightarrow}\varepsilon+s<\varepsilon+s_2$ contradicting $\varepsilon+s_2<\varepsilon+s$] and $\varepsilon+s\neq\min(\mathbb{Q}\backslash\alpha)$ hence $0<(\varepsilon+s)^{-1}\in\text{rep}(\alpha)$. As $s_1,s_3\in\alpha\Rightarrow s=\max(s_1,s_2)\Rightarrow s\in\alpha$ and from $0<s_1\leqslant\max(x_1, s_2)\Rightarrow 0<s$. The last two results gives us $s\cdot(\varepsilon+s)^{-1}\in\alpha\odot\text{rep}(\alpha)$. Now from $s_1\leqslant s$ we have by $0<x^{-1},(1-x)$ that $\varepsilon=s_1\cdot(1-x)\cdot x^{-1}\leqslant s\cdot(1-x)\cdot x^{-1}\Rightarrow 0<\varepsilon+s\leqslant s+s\cdot(1-x)\cdot x^{-1}=s\cdot(1+(1-x)\cdot x^{-1})\Rightarrow s^{-1}\cdot(1+(1-x)\cdot x^{-1})=(s\cdot(1+(1-x)\cdot x^{-1}))^{-1}\leqslant(\varepsilon+s)^{-1}\Rightarrow(1+(1-x)\cdot x^{-1})^{-1}\leqslant s\cdot(\varepsilon+s)^{-1}$. Now $1+(1-x)\cdot x^{-1}=1+x^{-1}-x\cdot x^{-1}=x^{-1}\Rightarrow x\leqslant s\cdot(\varepsilon+s)^{-1}$, as $s\cdot(\varepsilon+s)^{-1}\in\alpha\odot\alpha^{-1}$ we have by 8.3 that $x\in\alpha\odot\text{rep}(\alpha)$

So we have proved that $\alpha \odot \mathrm{rep}(\alpha) = \alpha_1$ and thus

$$\begin{aligned} \alpha \cdot \alpha^{-1} &= \alpha \odot \alpha^{-1} \\ &= \alpha \odot \mathrm{rep}(\alpha) \\ &= \alpha_1 \\ &\underset{\text{commutativity}}{=} \alpha^{-1} \cdot \alpha \end{aligned}$$

b. $(\boldsymbol{\alpha} \in \mathbb{R}_-)$ We have then

$$\begin{aligned} \alpha \cdot \alpha^{-1} &= \alpha \cdot (-(\mathrm{rep}(-\alpha))) \\ &= -(\alpha \cdot (-\alpha)^{-1}) \\ &= -(-((-\alpha) \cdot (-\alpha)^{-1})) \\ &= ((-\alpha) \cdot (-\alpha)^{-1}) \\ &\underset{3.b}{=} a_1 \\ &\underset{\text{commutativity}}{=} \alpha^{-1} \cdot \alpha \end{aligned}$$

4. **(Distributive)** We have the following cases to consider

a. $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_+$ We have then

$$x \in \alpha \cdot (\beta + \gamma) \;\Rightarrow\; x \in \alpha \odot (\beta + \gamma)$$

we have then two possible cases

i. $(\boldsymbol{x} \leqslant \mathbf{0})$ then $x \in \alpha \odot \beta$ also $0 \in \alpha \odot \gamma \Rightarrow x = x + 0 \in (\alpha \odot \beta) + (\alpha \odot \gamma)$

ii. $(\mathbf{0} < \boldsymbol{x})$ then $x = s \cdot t$ where $s \in \alpha \wedge 0 < s$ and $t \in \beta + \gamma \wedge 0 < t \Rightarrow \exists u \in \beta \wedge v \in \gamma$ with $t = u + v$. Using 7.10 we have that $x = s \cdot t = s \cdot (u + v) = s \cdot u + s \cdot v$. We have now the following possibilities

A. $u \leqslant 0 \wedge v \leqslant 0$ this case does not apply because it gives rise to the contradiction $0 < t = u + v \leqslant 0 \Rightarrow 0 < 0$.

B. $u \leqslant 0 \wedge 0 < v \underset{0<s}{\Rightarrow} s \cdot u \leqslant 0 \Rightarrow s \cdot u \in \alpha \odot \beta \wedge s \cdot t \in \alpha \odot \gamma \Rightarrow x = s \cdot u + s \cdot v \in (\alpha \odot \beta) + (\alpha \odot \gamma) \Rightarrow x \in (\alpha \odot \beta) + (\alpha \odot \gamma)$

C. $0 < u \wedge v \leqslant 0 \Rightarrow s \cdot v \leqslant 0 \Rightarrow s \cdot u \in \alpha \odot \beta \wedge s \cdot t \in \alpha \odot \gamma \Rightarrow x = s \cdot u + s \cdot v \in (\alpha \odot \beta) + (\alpha \odot \gamma) \Rightarrow x \in (\alpha \odot \beta) + (\alpha \odot \gamma)$

D. $0 < u \wedge 0 < v \Rightarrow s \cdot u \in \alpha \odot \beta \wedge s \cdot t \in \alpha \odot \gamma \Rightarrow x = s \cdot u + s \cdot v \in (\alpha \odot \beta) + (\alpha \odot \gamma) \Rightarrow x \in (\alpha \odot \beta) + (\alpha \odot \gamma)$

So we have $\alpha \odot (\beta + \gamma) \subseteq \alpha \odot \beta + \alpha \odot \gamma$

Consider now

$$x \in \alpha \odot \beta + \alpha \odot \gamma \quad \text{then}\, x = r + t \quad \text{with}\, r \in \alpha \odot \beta \wedge t \in \alpha \odot \gamma$$

we have then the following possibilities

i. $(\boldsymbol{x} \leqslant \mathbf{0}) \Rightarrow x \in \alpha \odot (\beta + \gamma)$

ii. $(\mathbf{0} < \boldsymbol{x})$ we have now the following possible sub-cases

A. $(\boldsymbol{r} \leqslant \mathbf{0} \wedge \boldsymbol{t} \leqslant \mathbf{0})$ then we would have the contradiction $x \leqslant 0$ so this case does not applies.

B. $(\boldsymbol{r} \leqslant \mathbf{0} \wedge \mathbf{0} < \boldsymbol{t})$ then $\exists u \in \alpha, \exists v \in \gamma$ such that $t = u \cdot v \wedge 0 < u \wedge 0 < v \Rightarrow t = u \cdot v = u \cdot (0+v) \in \alpha \odot (\beta + \gamma)$ since $r \leqslant 0 \Rightarrow x = r+t \leqslant 0+t \Rightarrow x \leqslant t$ we have by $t \in \alpha \odot (\beta+\gamma)$ and 8.3 that $x \in \alpha \odot (\beta + \gamma)$

C. $(\mathbf{0} < \boldsymbol{r} \wedge \boldsymbol{t} \leqslant \mathbf{0})$ then $\exists u \in \alpha, \exists v \in \gamma$ such that $r = u \cdot v \Rightarrow r = u \cdot v + u \cdot 0 = u \cdot (v+0) \in \alpha \odot (\beta + \gamma)$ since $t \leqslant 0 \Rightarrow x = r+t \leqslant r \Rightarrow x \leqslant r$ and from 8.3 and $r \in \alpha \odot (\beta+\gamma)$ we would have $x \in \alpha \odot (\beta+\gamma)$

D. $(\mathbf{0} < \boldsymbol{r} \wedge \mathbf{0} < \boldsymbol{t})$ then there exists $u \in \alpha, v \in \beta, u' \in \alpha, v' \in \gamma$ with $r = u \cdot v \wedge t = u' \cdot v' \wedge 0 < u \wedge 0 < v \wedge 0 < u' \wedge 0 < v'$. We have now to consider the following possibilities

1. $(\boldsymbol{u = u'})$ then $x = u \cdot v + u' \cdot v' = u \cdot v + u \cdot v' = u \cdot (v+v') \underset{0<u \wedge 0 < v+v'}{\Rightarrow} x = u \cdot (v+v') \in \alpha \odot (\beta+\gamma) \Rightarrow x \in \alpha \odot (\beta+\gamma)$

2. $(\boldsymbol{u < u'})$ then as $0 < u' \wedge 0 < v+v'$ we have $u' \cdot (v+v') \in \alpha \odot (\beta+\gamma)$ and from $u < u' \Rightarrow u \cdot v < u' \cdot v' \Rightarrow x = u \cdot v + u' \cdot v' \leqslant u' \cdot v + u' \cdot v' = u' \cdot (v+v') \in \alpha \odot (\beta+\gamma) \underset{8.3}{\Rightarrow} x \in \alpha \odot (\beta+\gamma)$

3. $(\boldsymbol{u' < u})$ then as $0 < u \wedge 0 < v+v'$ we have $u \cdot (v+v') \in \alpha \odot (\beta+\gamma)$. From $u' < u$ we have $u' \cdot v' < u \cdot .v' \Rightarrow x = u \cdot v + u' \cdot v' < u \cdot v + u \cdot v' = u \cdot (v+v') \in \alpha \odot (\beta+\gamma) \underset{8.3}{\Rightarrow} x \in \alpha \odot (\beta + \gamma)$

So we have $\alpha \odot \beta + \alpha \odot \gamma \subseteq \alpha \odot (\beta + \gamma) \Rightarrow \alpha \odot (\beta + \gamma) = \alpha \odot \beta + \alpha \odot \gamma$. So finally we have

$$\begin{aligned} \alpha \cdot (\beta+\gamma) &= \alpha \odot (\beta+\gamma) \\ &= \alpha \odot \beta + \alpha \odot \gamma \\ &= \alpha \cdot \beta + \alpha \cdot \gamma \end{aligned}$$

b. $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_-$ We have now two sub-cases

i. $(\boldsymbol{\beta+\gamma \in \mathbb{R}_+})$ First as $\gamma \in \mathbb{R}_-$ we have $-\gamma \in \mathbb{R}_+$ and thus we have using (4.a) that

$$\begin{aligned} \alpha \cdot (\beta+(-\gamma)) &= \alpha \cdot \beta + \alpha \cdot (-\gamma) \\ &= \alpha \cdot \beta + (-(\alpha \cdot \gamma)) \end{aligned}$$

now note that a $\beta, -\gamma \in \mathbb{R}_+$ we have $0 \in \beta, 0 \in -\gamma \Rightarrow 0 = 0+0 \in \beta+(-\gamma) \Rightarrow \beta+(-\gamma) \in \mathbb{R}_+$, next as we have that

$$\beta+\beta = (\beta+(-\gamma))+(\beta+\gamma)$$

we have then by multiplication both sides with α

$$\begin{array}{lcl}\alpha\cdot(\beta+\beta) & = & \alpha\cdot((\beta+(-\gamma))+(\beta+\gamma)) \\ & \underset{\beta\in\mathbb{R}_+,\beta+\gamma\in\mathbb{R}_+,\beta+\gamma\in\mathbb{R}_+\text{ and }4.a}{\Rightarrow} & \\ \alpha\cdot\beta+\alpha\cdot\beta & = & \alpha\cdot(\beta+(-\gamma))+\alpha\cdot(\beta+\gamma) \\ & \underset{\text{previous remark}}{=} & \alpha\cdot\beta+(-(\alpha\cdot\gamma))+\alpha\cdot(\beta+\gamma) \\ & \Rightarrow & \\ \alpha\cdot\beta+\alpha\cdot\gamma & = & \alpha\cdot(\beta+\gamma)\end{array}$$

ii. $\boldsymbol{(\beta+\gamma\in\mathbb{R}_-)}$ Then we have $-(\beta+\gamma)=(-\gamma)+(-\beta)\in\mathbb{R}_+$, $-\gamma\in\mathbb{R}_+, -\beta\in R_-$ and thus

$$\begin{array}{rcl}\alpha\cdot(\beta+\gamma) & = & -(\alpha\cdot(-(\beta+\gamma))) \\ & = & -(\alpha\cdot((-\beta)+(-\gamma))) \\ & = & -(\alpha\cdot((-\gamma)+(-\beta))) \\ & \underset{4.a.i}{=} & -(\alpha\cdot(-\gamma)+\alpha\cdot(-\beta)) \\ & = & -(-(\alpha\cdot\gamma)+(-(\alpha\cdot\beta))) \\ & = & \alpha\cdot\gamma+\alpha\cdot\beta \\ & = & \alpha\cdot\beta+\alpha\cdot\gamma\end{array}$$

c. $\alpha\in\mathbb{R}_+, \beta\in\mathbb{R}_+, \gamma=0$ Here we have

$$\begin{array}{rcl}\alpha\cdot(\beta+\gamma) & = & a\cdot(\beta+0) \\ & = & \alpha\cdot\beta \\ & = & \alpha\cdot\beta+0 \\ & = & \alpha\cdot\beta+\alpha\cdot 0 \\ & = & \alpha\cdot\beta+\alpha\cdot\gamma\end{array}$$

d. $\alpha\in\mathbb{R}_+, \beta\in\mathbb{R}_-, \gamma\in\mathbb{R}_+$ Here we have

$$\begin{array}{rcl}\alpha\cdot(\beta+\gamma) & = & \alpha\cdot(\gamma+\beta) \\ & \underset{4.b}{=} & \alpha\cdot\gamma+\alpha\cdot\beta \\ & = & \alpha\cdot\beta+\alpha\cdot\gamma\end{array}$$

e. $\alpha\in\mathbb{R}_+, \beta\in\mathbb{R}_-, \gamma\in\mathbb{R}_-$ then we have $-\beta, -\gamma\in\mathbb{R}_+$ and then we have

$$\begin{array}{rcl}\alpha\cdot(\beta+\gamma) & = & -(\alpha\cdot(-(\beta+\gamma))) \\ & = & -(\alpha\cdot((-\beta)+(-\gamma))) \\ & \underset{4.a}{=} & -(\alpha\cdot(-\beta)+\alpha\cdot(-\gamma)) \\ & = & -(-(\alpha\cdot\beta)+(-(\alpha\cdot\gamma))) \\ & = & \alpha\cdot\beta+\alpha\cdot\gamma\end{array}$$

f. $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_-, \gamma = 0$ We have then

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= \alpha \cdot \beta \\ &= \alpha \cdot \beta + 0 \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

g. $\alpha \in \mathbb{R}_+, \beta = 0, \gamma \in \mathbb{R}_+$ then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= \alpha \cdot \gamma \\ &= 0 + \alpha \cdot \gamma \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

h. $\alpha \in \mathbb{R}_+, \beta = 0, \gamma \in \mathbb{R}_-$ then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= \alpha \cdot \gamma \\ &= 0 + \alpha \cdot \gamma \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

i. $\alpha \in \mathbb{R}_+, \beta = 0, \gamma = 0$ then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= \alpha \cdot \gamma \\ &= 0 + \alpha \cdot \gamma \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

j. $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_+$ then we have $-\alpha \in \mathbb{R}_+$ and thus

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= -((-\alpha) \cdot (\beta + \gamma)) \\ &\underset{4.a}{=} -((-\alpha) \cdot \beta + (-\alpha) \cdot \gamma) \\ &= -(-(\alpha \cdot \beta) + (-(\alpha \cdot \gamma))) \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

k. $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_-$then we have $-\alpha \in \mathbb{R}_+$ and thus

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= -((-\alpha) \cdot (\beta + \gamma)) \\ &\underset{4.b}{=} -((-\alpha) \cdot \beta + (-\alpha) \cdot \gamma) \\ &= -(-(\alpha \cdot \beta) + (-(\alpha \cdot \gamma))) \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

l. $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_+, \gamma = 0$ then we have $-\alpha \in \mathbb{R}_+$ and thus

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= -((-\alpha) \cdot (\beta + \gamma)) \\ &\underset{4.c}{=} -((-\alpha) \cdot \beta + (-\alpha) \cdot \gamma) \\ &= -(-(\alpha \cdot \beta) + (-(\alpha \cdot \gamma))) \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

m. $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_+$ then we have $-\alpha \in \mathbb{R}_+$ and thus

$$\begin{array}{rcl}\alpha \cdot (\beta+\gamma) & = & -((-\alpha) \cdot (\beta+\gamma)) \\ & \underset{4.d}{=} & -((-\alpha) \cdot \beta + (-\alpha) \cdot \gamma) \\ & = & -(-(\alpha \cdot \beta)+(-(\alpha \cdot \gamma))) \\ & = & \alpha \cdot \beta+\alpha \cdot \gamma\end{array}$$

n. $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_-$ then we have $-\alpha \in \mathbb{R}_+$ and thus

$$\begin{array}{rcl}\alpha \cdot (\beta+\gamma) & = & -((-\alpha) \cdot (\beta+\gamma)) \\ & \underset{4.e}{=} & -((-\alpha) \cdot \beta + (-\alpha) \cdot \gamma) \\ & = & -(-(\alpha \cdot \beta)+(-(\alpha \cdot \gamma))) \\ & = & \alpha \cdot \beta+\alpha \cdot \gamma\end{array}$$

o. $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_-, \gamma = 0$ then we have $-\alpha \in \mathbb{R}_+$ and thus

$$\begin{array}{rcl}\alpha \cdot (\beta+\gamma) & = & -((-\alpha) \cdot (\beta+\gamma)) \\ & \underset{4.f}{=} & -((-\alpha) \cdot \beta + (-\alpha) \cdot \gamma) \\ & = & -(-(\alpha \cdot \beta)+(-(\alpha \cdot \gamma))) \\ & = & \alpha \cdot \beta+\alpha \cdot \gamma\end{array}$$

p. $\alpha \in \mathbb{R}_-, \beta = 0, \gamma \in \mathbb{R}_+$ then we have $-\alpha \in \mathbb{R}_+$ and thus

$$\begin{array}{rcl}\alpha \cdot (\beta+\gamma) & = & -((-\alpha) \cdot (\beta+\gamma)) \\ & \underset{4.g}{=} & -((-\alpha) \cdot \beta + (-\alpha) \cdot \gamma) \\ & = & -(-(\alpha \cdot \beta)+(-(\alpha \cdot \gamma))) \\ & = & \alpha \cdot \beta+\alpha \cdot \gamma\end{array}$$

q. $\alpha \in \mathbb{R}_-, \beta = 0, \gamma \in \mathbb{R}_-$ then we have $-\alpha \in \mathbb{R}_+$ and thus

$$\begin{array}{rcl}\alpha \cdot (\beta+\gamma) & = & -((-\alpha) \cdot (\beta+\gamma)) \\ & \underset{4.h}{=} & -((-\alpha) \cdot \beta + (-\alpha) \cdot \gamma) \\ & = & -(-(\alpha \cdot \beta)+(-(\alpha \cdot \gamma))) \\ & = & \alpha \cdot \beta+\alpha \cdot \gamma\end{array}$$

r. $\alpha \in \mathbb{R}_-, \beta = 0, \gamma = 0$ then we have $-\alpha \in \mathbb{R}_+$ and thus

$$\begin{array}{rcl}\alpha \cdot (\beta+\gamma) & = & -((-\alpha) \cdot (\beta+\gamma)) \\ & \underset{4.i}{=} & -((-\alpha) \cdot \beta + (-\alpha) \cdot \gamma) \\ & = & -(-(\alpha \cdot \beta)+(-(\alpha \cdot \gamma))) \\ & = & \alpha \cdot \beta+\alpha \cdot \gamma\end{array}$$

s. $\alpha \in 0, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_+$ then we have

$$\begin{array}{rcl}\alpha \cdot (\beta+\gamma) & = & 0 \\ & = & 0+0 \\ & = & \alpha \cdot \beta+\alpha \cdot \gamma\end{array}$$

t. $\alpha \in 0, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_-$ then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= 0\\ &= 0 + 0\\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

u. $\alpha \in 0, \beta \in \mathbb{R}_+, \gamma = 0$ then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= 0\\ &= 0 + 0\\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

v. $\alpha \in 0, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_+$ then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= 0\\ &= 0 + 0\\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

w. $\alpha \in 0, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_-$ then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= 0\\ &= 0 + 0\\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

x. $\alpha \in 0, \beta \in \mathbb{R}_-, \gamma = 0$ then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= 0\\ &= 0 + 0\\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

y. $\alpha \in 0, \beta = 0, \gamma \in \mathbb{R}_+$ then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= 0\\ &= 0 + 0\\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

z. $\alpha \in 0, \beta = 0, \gamma \in \mathbb{R}_-$ then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= 0\\ &= 0 + 0\\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

aa. $\alpha \in 0, \beta = 0, \gamma = 0$ then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= 0\\ &= 0 + 0\\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

5. **(Associative)**

   a. $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_+$ First take $x \in \alpha \cdot (\beta \cdot \gamma)$ then we have the following cases to consider

      i. $(\boldsymbol{x} \leqslant \mathbf{0})$ then $x \in \{r \in \mathbb{Q} | r \leqslant 0\} \Rightarrow x \in (\alpha \cdot \beta) \cdot \gamma$

      ii. $(\mathbf{0} < \boldsymbol{x})$ then there exists a $s \in \alpha, t \in (\beta \cdot \gamma)$ with $x = s \cdot t$ and $0 < s, t$, as $0 < t \in \beta \cdot \gamma$ we must have that there exists a $u \in \beta, v \in \gamma$ such that $t = u \cdot v$ and $0 < u, v$. So we have $x = s \cdot t = s \cdot (u \cdot v) \underset{7.10}{=} (s \cdot u) \cdot v$. But $s \cdot u \in \alpha \cdot \beta \Rightarrow x = (s \cdot u) \cdot v \in (\alpha \cdot \beta) \cdot \gamma \Rightarrow x \in (\alpha \cdot \beta) \cdot \gamma$

   it follows that $\alpha \cdot (\beta \cdot \gamma) \subseteq (\alpha \cdot \beta) \cdot \gamma$. Now if $x \in (\alpha \cdot \beta) \cdot \gamma \underset{\text{commutativity}}{=} \gamma \cdot (\alpha \cdot \beta) \underset{\text{commutativity}}{=} \gamma \cdot (\beta \cdot \alpha) \subseteq_{\text{we have just proved this}} (\gamma \cdot \beta) \cdot \alpha \underset{\text{commutativity twice}}{=} \alpha \cdot (\beta \cdot \gamma) \Rightarrow x \in \alpha \cdot (\beta \cdot \gamma)$ proving that $(\alpha \cdot \beta) \cdot \gamma \subseteq \alpha \cdot (\beta \cdot \gamma)$ and thus finally $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$

   b. $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_-$ in this case we have $-\gamma \in \mathbb{R}_+$

$$\begin{aligned}
\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (-(\beta \cdot (-\gamma))) \\
&= -(\alpha \cdot (\beta \cdot (-\gamma))) \\
&\underset{5.a}{=} -((\alpha \cdot \beta) \cdot (-\gamma)) \\
&= (\alpha \cdot \beta) \cdot \gamma
\end{aligned}$$

   c. $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_+, \gamma = 0$ in this case we have

$$\begin{aligned}
\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (\beta \cdot 0) \\
&= \alpha \cdot 0 \\
&= 0 \\
&= (\alpha \cdot \beta) \cdot 0 \\
&= (\alpha \cdot \beta) \cdot \gamma
\end{aligned}$$

   d. $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_+$ we have then $-\beta \in \mathbb{R}_+$ and

$$\begin{aligned}
\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (-((-\beta) \cdot \gamma)) \\
&= -(\alpha \cdot ((-\beta) \cdot \gamma)) \\
&\underset{5.a}{=} -((\alpha \cdot (-\beta)) \cdot \gamma) \\
&= -((-(\alpha \cdot \beta)) \cdot \gamma) \\
&= (\alpha \cdot \beta) \cdot \gamma
\end{aligned}$$

   e. $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_-$ we have then $-\beta, -\gamma \in \mathbb{R}_+$

$$\begin{aligned}
\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (-((-\beta) \cdot \gamma)) \\
&= \alpha \cdot ((-\beta) \cdot (-\gamma)) \\
&= (\alpha \cdot (-\beta)) \cdot (-\gamma) \\
&= -((\alpha \cdot (-\beta)) \cdot \gamma) \\
&= -((-(\alpha \cdot \beta)) \cdot \gamma) \\
&= (\alpha \cdot \beta) \cdot \gamma
\end{aligned}$$

f. $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_-, \gamma = 0$ here we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (\beta \cdot 0)\\ &= \alpha \cdot 0\\ &= 0\\ &= (\alpha \cdot \beta) \cdot 0\\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

g. $\alpha \in \mathbb{R}_+, \beta = 0, \gamma \in \mathbb{R}_+$ we have now

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (0 \cdot \gamma)\\ &= \alpha \cdot 0\\ &= 0\\ &= 0 \cdot \gamma\\ &= (\alpha \cdot 0) \cdot \gamma\\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

h. $\alpha \in \mathbb{R}_+, \beta = 0, \gamma \in \mathbb{R}_-$ we have now

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (0 \cdot \gamma)\\ &= \alpha \cdot 0\\ &= 0\\ &= 0 \cdot \gamma\\ &= (\alpha \cdot 0) \cdot \gamma\\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

i. $\alpha \in \mathbb{R}_+, \beta = 0, \gamma = 0$ we have now

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (0 \cdot \gamma)\\ &= \alpha \cdot 0\\ &= 0\\ &= 0 \cdot \gamma\\ &= (\alpha \cdot 0) \cdot \gamma\\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

j. $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_+$ here $-\alpha \in \mathbb{R}_+$ and

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= -((-\alpha) \cdot (\beta \cdot \gamma))\\ &\underset{5.a}{=} -(((-\alpha) \cdot \beta) \cdot \gamma)\\ &= -((-(\alpha \cdot \beta)) \cdot \gamma)\\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

k. $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_-$ here $-\alpha \in \mathbb{R}_+$ and

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= -((-\alpha) \cdot (\beta \cdot \gamma))\\ &\underset{5.b}{=} -(((-\alpha) \cdot \beta) \cdot \gamma)\\ &= -((-(\alpha \cdot \beta)) \cdot \gamma)\\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

l. $\alpha\in\mathbb{R}_-, \beta\in\mathbb{R}_+, \gamma=0$ here $-\alpha\in\mathbb{R}_+$ and

$$\begin{aligned}\alpha\cdot(\beta\cdot\gamma) &= -((-\alpha)\cdot(\beta\cdot\gamma))\\ &\underset{5.c}{=} -(((-\alpha)\cdot\beta)\cdot\gamma)\\ &= -((-(\alpha\cdot\beta))\cdot\gamma)\\ &= (\alpha\cdot\beta)\cdot\gamma\end{aligned}$$

m. $\alpha\in\mathbb{R}_-, \beta\in\mathbb{R}_-, \gamma\in\mathbb{R}_+$ here $-\alpha\in\mathbb{R}_+$ and

$$\begin{aligned}\alpha\cdot(\beta\cdot\gamma) &= -((-\alpha)\cdot(\beta\cdot\gamma))\\ &\underset{5.d}{=} -(((-\alpha)\cdot\beta)\cdot\gamma)\\ &= -((-(\alpha\cdot\beta))\cdot\gamma)\\ &= (\alpha\cdot\beta)\cdot\gamma\end{aligned}$$

n. $\alpha\in\mathbb{R}_-, \beta\in\mathbb{R}_-, \gamma\in\mathbb{R}_-$ here $-\alpha\in\mathbb{R}_+$ and

$$\begin{aligned}\alpha\cdot(\beta\cdot\gamma) &= -((-\alpha)\cdot(\beta\cdot\gamma))\\ &\underset{5.e}{=} -(((-\alpha)\cdot\beta)\cdot\gamma)\\ &= -((-(\alpha\cdot\beta))\cdot\gamma)\\ &= (\alpha\cdot\beta)\cdot\gamma\end{aligned}$$

o. $\alpha\in\mathbb{R}_-, \beta\in\mathbb{R}_-, \gamma=0$ here $-\alpha\in\mathbb{R}_+$ and

$$\begin{aligned}\alpha\cdot(\beta\cdot\gamma) &= -((-\alpha)\cdot(\beta\cdot\gamma))\\ &\underset{5.f}{=} -(((-\alpha)\cdot\beta)\cdot\gamma)\\ &= -((-(\alpha\cdot\beta))\cdot\gamma)\\ &= (\alpha\cdot\beta)\cdot\gamma\end{aligned}$$

p. $\alpha\in\mathbb{R}_-, \beta=0, \gamma\in\mathbb{R}_+$ here $-\alpha\in\mathbb{R}_+$ and

$$\begin{aligned}\alpha\cdot(\beta\cdot\gamma) &= -((-\alpha)\cdot(\beta\cdot\gamma))\\ &\underset{5.g}{=} -(((-\alpha)\cdot\beta)\cdot\gamma)\\ &= -((-(\alpha\cdot\beta))\cdot\gamma)\\ &= (\alpha\cdot\beta)\cdot\gamma\end{aligned}$$

q. $\alpha\in\mathbb{R}_-, \beta=0, \gamma\in\mathbb{R}_-$ here $-\alpha\in\mathbb{R}_+$ and

$$\begin{aligned}\alpha\cdot(\beta\cdot\gamma) &= -((-\alpha)\cdot(\beta\cdot\gamma))\\ &\underset{5.h}{=} -(((-\alpha)\cdot\beta)\cdot\gamma)\\ &= -((-(\alpha\cdot\beta))\cdot\gamma)\\ &= (\alpha\cdot\beta)\cdot\gamma\end{aligned}$$

r. $\alpha\in\mathbb{R}_-, \beta=0, \gamma=0$ here $-\alpha\in\mathbb{R}_+$ and

$$\begin{aligned}\alpha\cdot(\beta\cdot\gamma) &= -((-\alpha)\cdot(\beta\cdot\gamma))\\ &\underset{5.i}{=} -(((-\alpha)\cdot\beta)\cdot\gamma)\\ &= -((-(\alpha\cdot\beta))\cdot\gamma)\\ &= (\alpha\cdot\beta)\cdot\gamma\end{aligned}$$

s. $\alpha \in 0, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_+$ in this case we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma)\\ &= 0\\ &= 0 \cdot \gamma\\ &= (0 \cdot \beta) \cdot \gamma\\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

t. $\alpha \in 0, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_-$ in this case we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma)\\ &= 0\\ &= 0 \cdot \gamma\\ &= (0 \cdot \beta) \cdot \gamma\\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

u. $\alpha \in 0, \beta \in \mathbb{R}_+, \gamma = 0$ in this case we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma)\\ &= 0\\ &= 0 \cdot \gamma\\ &= (0 \cdot \beta) \cdot \gamma\\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

v. $\alpha \in 0, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_+$ in this case we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma)\\ &= 0\\ &= 0 \cdot \gamma\\ &= (0 \cdot \beta) \cdot \gamma\\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

w. $\alpha \in 0, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_-$ in this case we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma)\\ &= 0\\ &= 0 \cdot \gamma\\ &= (0 \cdot \beta) \cdot \gamma\\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

x. $\alpha \in 0, \beta \in \mathbb{R}_-, \gamma = 0$ in this case we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma)\\ &= 0\\ &= 0 \cdot \gamma\\ &= (0 \cdot \beta) \cdot \gamma\\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

y. $\alpha \in 0, \beta = 0, \gamma \in \mathbb{R}_+$ in this case we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma)\\ &= 0\\ &= 0 \cdot \gamma\\ &= (0 \cdot \beta) \cdot \gamma\\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

z. $\alpha \in 0, \beta = 0, \gamma \in \mathbb{R}_-$ in this case we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma)\\ &= 0\\ &= 0 \cdot \gamma\\ &= (0 \cdot \beta) \cdot \gamma\\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

aa. $\alpha \in 0, \beta = 0, \gamma = 0$ in this case we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma)\\ &= 0\\ &= 0 \cdot \gamma\\ &= (0 \cdot \beta) \cdot \gamma\\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

$\square$

Notation 8.26. *If $x, y \in \mathbb{R}, x \neq 0$ then we note*

1. $x^{-1} \underset{\text{notation}}{=} \frac{1}{x}$
2. $y \cdot x^{-1} = \frac{y}{x}$

Theorem 8.27. $\forall r, s \in \mathbb{Q}$ *we have*

1. $\alpha_r + \alpha_s = \alpha_{r+s}$
2. $\alpha_r \cdot \alpha_s = \alpha_{r \cdot s}$
3. *if* $\alpha_r \neq 0$ *then* $(\alpha_r)^{-1} = \alpha_{r^{-1}}$

Proof.

1. If $x \in \alpha_r + \alpha_s$ then there exists $u \in \alpha_r, v \in \alpha_s$ with $x = u + v$. From $u \in \alpha_r$, $v \in \alpha_s$ we have $u < r, v < s \Rightarrow u + v < r + v, r + v < r + s \Rightarrow u + v < r + s \Rightarrow x = v + v \in \alpha_{r+s} \Rightarrow \alpha_r + \alpha_s \subseteq \alpha_{r+s}$. If $x \in \alpha_{r+s} \Rightarrow x < r + s \Rightarrow x - r < s$. Using the density theorem of $\mathbb{Q}$ (see 7.36) there exists a $z \in \mathbb{Q}$ with $x - r < z < s$. Then $z \in \alpha_s$ and $\varepsilon = z - (x - r) > 0 \Rightarrow -\varepsilon < 0 \Rightarrow r + (-\varepsilon) < r \Rightarrow r - \varepsilon \in \alpha_r$. As $\alpha_r + \alpha_s \ni r - \varepsilon + z = r - (z - (x - r)) + z = r - z + x - r + z = x \Rightarrow x \in \alpha_r + \alpha_s \Rightarrow \alpha_{r+s} \subseteq \alpha_r + \alpha_s$. So we conclude that $\alpha_r + \alpha_s = \alpha_{r+s}$.

2. We have the following possible cases for α_r, α_s (see 8.15)

   a. $(\boldsymbol{\alpha_r \in \mathbb{R}_+, \alpha_s \in \mathbb{R}_+})$ First as $0 \in \alpha_r, 0 \in \alpha_s \Rightarrow 0 < r \wedge 0 < s$. Take $x \in \alpha_r \odot \alpha_s$ then we have the following possibilities:

      i. $(\boldsymbol{x \leqslant 0})$ as $0 < r, 0 < s$ we have $0 < r \cdot s \Rightarrow x \leqslant 0 < r \cdot s \Rightarrow x \in \alpha_{r \cdot s}$

      ii. $(\boldsymbol{0 < x})$ in this case we have $x = u \cdot v$ where $u \in \alpha_r, v \in \alpha_s$, $0 < u, 0 < v$. As $0 < u < r, 0 < v < s \Rightarrow u \cdot v < r \cdot v, r \cdot v < r \cdot s \Rightarrow u \cdot v < r \cdot s \Rightarrow x < r \cdot s \Rightarrow x \in \alpha_{r \cdot s}$

   we conclude thus that $\alpha_r \odot \alpha_s \subseteq \alpha_{r \cdot s}$. Take now $x \in \alpha_{r \cdot s} \Rightarrow x < r \cdot s$ we have then the following cases for x

      i. $(\boldsymbol{x \leqslant 0})$ then trivially $x \in \alpha_r \odot \alpha_s$

      ii. $(\boldsymbol{0 < x})$ then $0 < x$. As $0 < r$ we have by the density theorem (see 7.36) that there exists a $\varepsilon_1 \in \mathbb{Q}$ such that $0 < \varepsilon_1 < r$. From $x < r \cdot s$ we have $0 < r \cdot s - x \underset{0 < s \Rightarrow 0 < s^{-1}}{\Rightarrow} 0 < (r \cdot s - x) \cdot s^{-1}$ and by the density theorem again we have that there exists a $\varepsilon_2 \in \mathbb{Q}$ such that $0 < \varepsilon_2 < (r \cdot s - x) \cdot s^{-1}$. Take now $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$ then $0 < \varepsilon \leqslant \varepsilon_1 < r$ and $0 < \varepsilon \leqslant \varepsilon_2 < (r \cdot s - x) \cdot s^{-1} \Rightarrow -((r \cdot s - x) \cdot s^{-1}) < -\varepsilon \Rightarrow r - ((r \cdot s - x) \cdot s^{-1}) < r - \varepsilon \Rightarrow x \cdot s^{-1} < r - \varepsilon \underset{0 < x, 0 < s^{-1} \Rightarrow 0 \in x \cdot s^{-1}}{\Rightarrow} 0 < x \cdot s^{-1} < r - \varepsilon \underset{7.35}{\Rightarrow} 0 < (r - \varepsilon)^{-1} < x^{-1} \cdot s \underset{0 < x}{\Rightarrow} 0 < x \cdot (r - \varepsilon)^{-1} < s \Rightarrow 0 < x \cdot (r - \varepsilon)^{-1} \in \alpha_s$. As $\varepsilon < 0 \Rightarrow 0 < -\varepsilon$ and $\varepsilon < r$ we have $0 < r - \varepsilon = r + (-\varepsilon) < r \Rightarrow 0 < r - \varepsilon \in \alpha_r$ and thus $(r - \varepsilon)(x \cdot (r - \varepsilon)^{-1}) \in \alpha_r \odot \alpha_s \Rightarrow x \in \alpha_r \odot \alpha_s$. We conclude thus that $\alpha_{r \cdot s} \subseteq \alpha_r \odot \alpha_s$.

   So we finally reach the conclusion that $\alpha_{r \cdot s} = \alpha_r \odot \alpha_s$. As $\alpha_r, \alpha_s \in \mathbb{R}_+$ we have $\alpha_r \cdot \alpha_s = \alpha_r \odot \alpha_s = \alpha_{r \cdot s} \Rightarrow \alpha_r \cdot \alpha_s = \alpha_{r \cdot s}$

   b. $(\boldsymbol{\alpha_r \in \mathbb{R}_+, \alpha_s \in \mathbb{R}_-})$ Then $\alpha_{-s} \underset{8.9}{=} -\alpha_s \in \mathbb{R}_+$ and $\alpha_r \cdot \alpha_s = \alpha_r \cdot (-(-\alpha_s)) = -(\alpha_r \cdot (-\alpha_s)) = -(\alpha_r \cdot \alpha_{-s}) \underset{(a)}{=} -\alpha_{r \cdot (-s)} = -\alpha_{-(r \cdot s)} \underset{8.9}{=} -(-(\alpha_{r \cdot s})) = \alpha_{r \cdot s}$

   c. $(\boldsymbol{\alpha_r \in \mathbb{R}_+, \alpha_s = 0}) \Rightarrow \alpha_s = \alpha_0 \underset{8.4}{\Rightarrow} s = 0$. Then $\alpha_r \cdot \alpha_s = \alpha_r \cdot 0 = 0 = \alpha_0 = \alpha_{r \cdot 0} = \alpha_{r \cdot s}$

   d. $(\boldsymbol{\alpha_r \in \mathbb{R}_-, \alpha_s \in \mathbb{R}_+})$ Then $\alpha_r \cdot \alpha_s = \alpha_s \cdot \alpha_r \underset{(b)}{=} \alpha_{s \cdot r} = \alpha_{r \cdot s}$

   e. $(\boldsymbol{\alpha_r \in \mathbb{R}_-, \alpha_s \in \mathbb{R}_-})$ Then $\alpha_{-r} = -\alpha_r \in \mathbb{R}_+, \alpha_{-s} = -\alpha_s \in \mathbb{R}_+$ and $\alpha_r \cdot \alpha_s = (-(-\alpha_r)) \cdot (-(-\alpha_s)) = (-\alpha_r) \cdot (-\alpha_s) = \alpha_{-r} \cdot \alpha_s \underset{(a)}{=} \alpha_{(-r) \cdot (-s)} = \alpha_{r \cdot s}$

   f. $(\boldsymbol{\alpha_r \in \mathbb{R}_-, \alpha_s = 0}) \Rightarrow \alpha_s = \alpha_0 \underset{8.4}{\Rightarrow} s = 0$. Then $\alpha_r \cdot \alpha_s = \alpha_r \cdot 0 = 0 = \alpha_0 = \alpha_{r \cdot 0} = \alpha_{r \cdot s}$

   g. $(\boldsymbol{\alpha_r = 0, \alpha_s \in \mathbb{R}_+}) \Rightarrow \alpha_r = \alpha_0 \Rightarrow r = 0$. Then $\alpha_r \cdot \alpha_s = 0 \cdot \alpha_s = 0 = \alpha_0 = \alpha_{0 \cdot s} = \alpha_{r \cdot s}$

   h. $(\boldsymbol{\alpha_r = 0, \alpha_s \in \mathbb{R}_-}) \Rightarrow \alpha_r = \alpha_0 \Rightarrow r = 0$. Then $\alpha_r \cdot \alpha_s = 0 \cdot \alpha_s = 0 = \alpha_0 = \alpha_{0 \cdot s} = \alpha_{r \cdot s}$

   i. $(\boldsymbol{\alpha_r = 0, \alpha_s = 0}) \Rightarrow \alpha_r = \alpha_0 \Rightarrow r = 0$. Then $\alpha_r \cdot \alpha_s = 0 \cdot \alpha_s = 0 = \alpha_0 = \alpha_{0 \cdot s} = \alpha_{r \cdot s}$

3. This follows from 8.24. □

Theorem 8.28. *$\langle \mathbb{Q}_{\mathbb{R}}, +, \cdot \rangle$ forms a sub-field of $\langle \mathbb{R}, +, \cdot \rangle$, further the function $i_{\mathbb{Q}}$: $\mathbb{Q} \to \mathbb{Q}_{\mathbb{R}}$ defined by $r \to \alpha_r$ is a field isomorphism. So we can consider $\mathbb{Q}_{\mathbb{R}}$ as the set of rational numbers embedded in the set of reals.*

Proof. First we prove that $\langle \mathbb{Q}_{\mathbb{R}}, +, \cdot \rangle$ is a sub field

1. If $x, y \in \mathbb{Q}_{\mathbb{R}} \Rightarrow \exists r, s \in \mathbb{Q} \vdash x = \alpha_r$ and $y = \alpha_s$ and thus $x + y = \alpha_r + \alpha_s \underset{\text{previous theorem}}{=} \alpha_{r+s} \in \mathbb{Q}_{\mathbb{R}}$
2. If $x, y \in \mathbb{Q}_{\mathbb{R}} \Rightarrow \exists r, s \in \mathbb{Q} \vdash x = \alpha_r$ and $y = \alpha_s$ and thus $x \cdot y = \alpha_r \cdot \alpha_s \underset{\text{previous theorem}}{=} \alpha_{r \cdot s} \in \mathbb{Q}_{\mathbb{R}}$
3. If $x \in \mathbb{Q}_{\mathbb{R}} \backslash \{0\}$ then $\exists r \in \mathbb{Q}$ with $x = \alpha_r \neq 0 = \alpha_0 \Rightarrow r \in \mathbb{Q} \backslash \{0\} \Rightarrow r^{-1}$ exists and if we take $x' = \alpha_{r^{-1}}$ then we have $x' \cdot x \underset{\text{commutativity}}{=} x \cdot x' = \alpha_r \cdot \alpha_{r^{-1}} = \alpha_{r \cdot r^{-1}} = \alpha_1 = 1$, so $x^{-1} = x' = \alpha_{r^{-1}} \in \mathbb{Q}_{\mathbb{R}}$
4. $0 = \alpha_0 \in \mathbb{Q}_{\mathbb{R}}$
5. $1 = \alpha_1 \in \mathbb{Q}_{\mathbb{R}}$

Next we prove that $i_{\mathbb{Q}}$ is a bijection

1. **(injectivity)** If $r, s \in \mathbb{Q}$ is such that $i_{\mathbb{Q}}(r) = i_{\mathbb{Q}}(s) = \alpha_r = \alpha_s \underset{8.4}{\Rightarrow} r = s$
2. **(surjectivity)** If $x \in \mathbb{Q}_{\mathbb{R}} \Rightarrow \exists r \in \mathbb{Q} \vdash x = \alpha_r = i_{\mathbb{Q}}(r)$.

Finally we prove the homeomorphism properties

1. If $r, s \in \mathbb{Q}_{\mathbb{R}} \Rightarrow i_{\mathbb{Q}}(r) + i_{\mathbb{Q}}(s) = \alpha_r + \alpha_s \underset{\text{previoustheorem}}{=} \alpha_{r+s} = i_{\mathbb{Q}}(r+s)$
2. If $r, s \in \mathbb{Q}_{\mathbb{R}} \Rightarrow i_Q(r \cdot s) = \alpha_{r \cdot s} \underset{\text{previoustheorem}}{=} \alpha_r \cdot \alpha_s = i_{\mathbb{Q}}(r) \cdot i_{\mathbb{Q}}(s)$
3. $i_{\mathbb{Q}}(1) = \alpha_1 = 1$ □

Corollary 8.29. *The set $\mathbb{Q}_{\mathbb{R}}$ is denumerable*

Proof. As $\mathbb{Q}$ is denumerable (see 7.40) there exists a bijection β: $\mathbb{N}_0 \to \mathbb{Q}$, as by the previous theorem we have that $i_{\mathbb{Q}}$: $\mathbb{Q} \to \mathbb{Q}_{\mathbb{R}}$ is a bijection we have that $i_{\mathbb{Q}} \circ \beta$: $\mathbb{N}_0 \to \mathbb{Q}_{\mathbb{R}}$ is a bijection proving that $\mathbb{Q}_{\mathbb{R}}$ is denumerable. □

8.2.3 Power in $\mathbb{R}$

Definition 8.30. *As $\langle \mathbb{R}, \cdot \rangle$ is a abelian semi-group we have by 4.22 that given a $a \in \mathbb{R}$ and $n \in \mathbb{N}_0$ that there exists a a^n such that*

$$\begin{aligned} a^0 &= 1 \\ a^{n+1} &= a^n \cdot a \underset{\text{abelian}}{=} a \cdot a^n \end{aligned}$$

Theorem 8.31. *If $n, n' \in \mathbb{N}_0$ and $a \in \mathbb{R}$ then $a^{n'+n} = a^{n'} \cdot a^n$*

Proof. We prove this by induction on n. So let $X = \{n \in \mathbb{N}_0 | a^{n'+n} = a^{n'} \cdot a^n\}$ then we have

1. If $n = 0$ then $a^{n'+n} = a^{n'+0} = a^{n'} = a^{n'} \cdot 1 = a^{n'} \cdot a^0 \Rightarrow 0 \in X$

2. If $n \in X$ then $a^{n'+(n+1)} = a^{(n'+n)+1} = a^{(n'+n)} \cdot a \underset{n \in X}{=} (a^{n'} \cdot a^n) \cdot a = a^{n'} \cdot (a^n \cdot a) = a^{n'} \cdot a^{n+1}$ and thus $n+1 \in X$

Using mathematical induction (see 4.10) we have $X = \mathbb{N}_0$ proving the theorem □

Theorem 8.32. *If $n \in \mathbb{N}_0$ and $a, b \in \mathbb{R}$ then we have $(a \cdot b)^n = a^n \cdot b^n$*

Proof. We prove this by induction so take $\mathcal{S} = \{n \in \mathbb{N}_0 | (a \cdot b)^n = a^n \cdot b^n\}$ then we have

$\boldsymbol{n = 0}$**.** then $(a \cdot b)^0 = 1 = 1 \cdot 1 = a^0 \cdot b^0$

$\boldsymbol{n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}}$**.** then $(a \cdot b)^{n+1} = (a \cdot b)^n \cdot (a \cdot b) = (a^n \cdot b^n) \cdot (a^n \cdot a) \cdot (b^n \cdot b) = a^{n+1} \cdot b^{n+1}$ □

Theorem 8.33. *In $\mathbb{R}$ we have*

$$\begin{aligned} 0^n &= 0\,(\text{if } n \neq 0) \\ 1^n &= 1 \\ (-1)^n &= -1 \text{ or } 1 \\ (-1)^{2 \cdot n} &= 1 \\ (-1)^{2 \cdot n+1} &= -1 \end{aligned}$$

Proof.

1. If $n \neq 0 \Rightarrow \exists m \in \mathbb{N}_0 \vdash n = m+1$ then $0^n = 0^{(m+1)} = 0^m \cdot 0 = 0$
2. $1^n = 1$ is proved by induction on n, let $X = \{n \in \mathbb{N}_0 | 1^n = 1\}$ then
   a. $1^0 = 1 \Rightarrow 0 \in X$
   b. If $n \in X \Rightarrow 1^{n+1} = 1^n \cdot 1 \underset{n \in X}{=} 1 \cdot 1 = 1 \Rightarrow n+1 \in X$

   so $X = \mathbb{N}_0$
3. $(-1)^n = \pm 1$ is proved by induction on n, let $X = \{n \in \mathbb{N}_0 | (-1)^n = -1 \text{ or } 1\}$ then
   a. $(-1)^0 = 1 \Rightarrow 0 \in X$
   b. If $n \in X$ then $(-1)^{n+1} = (-1)^n \cdot (-1) \underset{n \in X}{=} (-1) \cdot (-1) \vee 1 \cdot (-1) = 1 \vee -1 \Rightarrow n+1 \in X$

   so $X = \mathbb{N}_0$
4. $(-1)^{2 \cdot n} = (-1)^{(1+1) \cdot n} = (-1)^{n+n} = (-1)^n \cdot (-1)^n \underset{(3)}{=} (-1) \cdot (-1) \text{ or } 1 \cdot 1 = 1$
5. $(-1)^{2 \cdot n+1} = (-1)^{2 \cdot n} \cdot (-1) \underset{(4)}{=} 1 \cdot (-1) = -1$ □

8.3 Order relation on $\mathbb{R}$

Theorem 8.34. *If $\alpha, \beta \in \mathbb{R}_+$ then $\alpha+\beta \in \mathbb{R}_+$ and $\alpha \cdot \beta \in \mathbb{R}_+$*

Proof. If $\alpha, \beta \in \mathbb{R}$ then $0 \in \alpha \wedge 0 \in \beta$ then $0 = 0+0 \in \alpha+\beta \Rightarrow \alpha+\beta \in \mathbb{R}_+$. Also $0 \in \{r \in \mathbb{Q} | r \leqslant 0\} \subseteq \alpha \odot \beta \underset{a,b \in \mathbb{R}_+}{=} \alpha \cdot \beta \Rightarrow \alpha \cdot \beta \in \mathbb{R}_+$ □

We define now relations $<\subseteq\mathbb{R}\times\mathbb{R}$ and $\leqslant\subseteq\mathbb{R}\times\mathbb{R}$ as follows

Definition 8.35. $<=\{(a,\beta)\in\mathbb{R}\times\mathbb{R}|\beta+(-\alpha)\in\mathbb{R}_+\}$, *so* $\alpha<\beta\Leftrightarrow(a,\beta)\in<\Leftrightarrow\beta+(-\alpha)=\beta-\alpha\in\mathbb{R}_+$.

Definition 8.36. $\leqslant=\{(\alpha,\beta)\in\mathbb{R}\times\mathbb{R}|\alpha=\beta\vee\beta+(-\alpha)\in\mathbb{R}_+\}$, *so* $\alpha\leqslant b\Leftrightarrow(\alpha,\beta)\in\leqslant\Leftrightarrow\alpha=\beta\vee\beta+(-\alpha)\in\mathbb{R}_+\Leftrightarrow\alpha=\beta\vee\alpha<\beta$

Theorem 8.37. $\forall\alpha,\beta\in\mathbb{R}$ *we have*

1. $\alpha<\beta$ *iff* $\alpha\subset\beta$
2. $\alpha\leqslant\beta$ *iff* $\alpha\subseteq\beta$

Proof. Be care full in this prove $<$ can mean a relation in $\mathbb{R}$ or in $\mathbb{Q}$, context tells us which is which.

1.
   a. **($\boldsymbol{\alpha<\beta\Rightarrow\alpha\subset\beta}$)** If $\alpha<\beta\Rightarrow\beta+(-\alpha)\in\mathbb{R}_+\Rightarrow0\in\beta+(-\alpha)$. If $r\in\alpha\underset{8.8}{\Rightarrow}-r\notin-\alpha\Rightarrow-r\in\mathbb{Q}\backslash(-\alpha)$, now as $0\in\beta+(-\alpha)$ there exists a $s\in\beta$, $t\in-\alpha$ such that $0=s+t\Rightarrow s=-t$. By 8.1, $t\in-\alpha$ and $-r\in\mathbb{Q}\backslash(-\alpha)$ we have $t<(-r)\underset{7.22}{\Rightarrow}r<(-t)=s\in\beta\underset{8.3}{\Rightarrow}r\in\beta$ and thus we have $\alpha\subseteq\beta$. If now $\alpha=\beta\Rightarrow\beta+(-\alpha)=0=\alpha_0\Rightarrow0\in\beta+(-\alpha)=\alpha_0\Rightarrow0<0$ a contradiction, so we must have $\alpha\neq\beta\Rightarrow\alpha\subset\beta$.

   b. **($\boldsymbol{\alpha\subset\beta\Rightarrow\alpha<\beta}$)** If $\alpha\subset\beta\Rightarrow\exists r\in\beta$ with $r\notin\alpha\Rightarrow r\in\mathbb{Q}\backslash\alpha$ as $\max(\beta)$ does not exists there exists a $r'\in\beta$with $r<r'$. If $r'\in\alpha$ then using 8.1 we have $r'<r$ contradicting $r<r'$ and thus we must have $r'\notin\alpha\Rightarrow r'\in\mathbb{Q}\backslash\alpha$. Hence $r,r'\in\mathbb{Q}\backslash\alpha$ and $r<r'\Rightarrow r'\neq\min(\mathbb{Q}\backslash\alpha)\Rightarrow-r'\in-\alpha$. We have then as $r'\in\beta$,$-r'\in-\alpha$ and $0=r'+(-r')\in\beta+(-\alpha)\Rightarrow\beta+(-\alpha)\in\mathbb{R}_+\Rightarrow\alpha<\beta$

2.
   a. **($\boldsymbol{\alpha\leqslant\beta\Rightarrow\alpha\subseteq\beta}$)** As $a\leqslant b$ we have the following two cases

      i. **($\boldsymbol{\alpha<\beta}$)** Using 1.a we have then $\alpha\subset\beta\Rightarrow\alpha\subseteq\beta$

      ii. **($\boldsymbol{\alpha=\beta}$)** $\Rightarrow\alpha\subseteq\beta$

   b. **($\boldsymbol{\alpha\subseteq\beta\Rightarrow\alpha\leqslant\beta}$)** As $\alpha\subseteq\beta$ we have the following two exclusive cases to consider

      i. **($\boldsymbol{\alpha=\beta}$)** then $\alpha=\beta\Rightarrow\alpha\leqslant\beta$

      ii. **($\boldsymbol{\alpha\subset\beta}$)** then using 1.b we have $\alpha<\beta\Rightarrow\alpha\leqslant\beta$ □

Theorem 8.38. $\langle\mathbb{R},\leqslant\rangle$ *forms a partially ordered set that is fully-ordered*

Proof.

1. **(Reflexivity)** If $\alpha\in\mathbb{R}\Rightarrow\alpha\subseteq\alpha\underset{8.37}{\Rightarrow}\alpha\leqslant\alpha$
2. **(Anti-symmetry)** If $\alpha,\beta\in\mathbb{R}$ with $\alpha\leqslant\beta$ and $\beta\leqslant\alpha$ then by 8.37 we have $\alpha\subseteq\beta$ and $\beta\subseteq\alpha$ so $\alpha=\beta$

3. **(Transitivity)** If $\alpha, \beta, \gamma \in \mathbb{R}$ with $\alpha \leqslant \beta$ and $\beta \leqslant \gamma$ then by 8.37 we have $\alpha \subseteq \beta$ and $\beta \subseteq \gamma \Rightarrow \alpha \subseteq \gamma$ and by 8.37 again we have $\alpha \leqslant \gamma$
4. **(Fully-ordered)** If $\alpha, \beta \in \mathbb{R}$ then by 8.15 we have for $\alpha + (-\beta)$ either
   a. $(\boldsymbol{\alpha} + (-\boldsymbol{\beta}) = \mathbf{0})$ From this we have $\alpha = \beta \Rightarrow \alpha \subseteq \beta \underset{8.37}{\Rightarrow} \alpha \leqslant \beta$
   b. $(\boldsymbol{\alpha} + (-\boldsymbol{\beta}) \in \mathbb{R}_+)$ From this we have by definition $\beta \leqslant \alpha$
   c. $(\boldsymbol{\alpha} + (-\boldsymbol{\beta}) \in \mathbb{R}_-)$ In this case we have $-(\alpha + (-\beta)) \in \mathbb{R}_+ \Rightarrow (-\alpha) + \beta \in \mathbb{R}_+ \Rightarrow \beta + (-\alpha) \in \mathbb{R}_+$ and thus $\alpha \leqslant \beta$ □

Theorem 8.39. *We have the following for the set of reals*

1. $\alpha \in \mathbb{R}_+$ *iff* $0 < \alpha$
2. $\alpha \in \mathbb{R}_-$ *iff* $\alpha < 0$
3. *If* $\alpha, \beta \in \mathbb{R}$ *then we have the following exclusive possibilities*
   a. $\alpha < \beta$
   b. $\beta < \alpha$
   c. $\alpha = \beta$
4. *If* $\alpha, \beta \in \mathbb{R}$ *with* $\alpha < \beta \Rightarrow -\beta < -\alpha$
5. *If* $\alpha, \beta, \gamma \in \mathbb{R}$ *with* $\alpha < \beta \Rightarrow \alpha + \gamma < \beta + \gamma$
6. *If* $\alpha, \beta \in \mathbb{R}$ *and* $\alpha < \beta$ *with* $\gamma \in \mathbb{R}_+$ *then* $\alpha \cdot \gamma < \beta \cdot \gamma$ *[as we have* $\alpha = \beta \Rightarrow \alpha \cdot \lambda = \beta \cdot \lambda$ *we have also that* $\alpha \leqslant \beta$ *implies* $\alpha \cdot \lambda \leqslant \beta \cdot \lambda$*]*
7. *If* $\alpha, \beta \in \mathbb{R}$ *and* $\alpha < \beta$ *with* $\gamma \in \mathbb{R}_-$ *then* $\beta \cdot \gamma < \alpha \cdot \gamma$ *[as we have* $\alpha = \beta \Rightarrow \alpha \cdot \lambda = \beta \cdot \lambda$ *we have also that* $\alpha \leqslant \beta$ *implies* $\beta \cdot \lambda \leqslant \alpha \cdot \lambda$*]*
8. *If* $0 < x \Rightarrow 0 < x^{-1}$
9. *If* $0 < x < y \Rightarrow y^{-1} < x^{-1}$
10. *If* $x \in \mathbb{R} \Rightarrow 0 \leqslant x^2$ *and if* $x \neq 0$ *then* $0 < x^2$
11. *If* $x, y \in \mathbb{R}_+ \bigcup \{0\}$ *are such that* $x < y$ *then* $x^2 < y^2$ *(so* $^2: \{x \in \mathbb{R} | 0 \leqslant x\} \rightarrow \{x \in \mathbb{R} | 0 \leqslant x\}$ *is a strictly increasing function)*
12. *If* $\alpha \in \mathbb{R}_+$ *and* $n \in \mathbb{N}_0$ *then* $\alpha^n \in \mathbb{R}_+$
13. *If* $\alpha \in \mathbb{R}_+$ *so that* $0 < \alpha < 1$ *then if* $n \in \mathbb{N}$ *we have* $0 < \alpha^n < \alpha$
14. *If* $\alpha \in \mathbb{R}$ *such that* $1 \leqslant \alpha$
    a. *if* $\alpha < \beta$ *then* $\alpha < \beta^n$
    b. *if* $\alpha \leqslant \beta$ *then* $\alpha \leqslant \beta^n$

Proof.

1. *If* $\alpha \in \mathbb{R}_+$ *then* $\alpha + (-0) = \alpha \in \mathbb{R}_+ \Rightarrow 0 < \alpha$. *On the other hand if* $0 < \alpha$ *then* $\mathbb{R}_+ \ni \alpha + (-0) = \alpha \Rightarrow \alpha \in \mathbb{R}_+$.
2. *If* $\alpha \in \mathbb{R}_-$ *then* $-\alpha \in \mathbb{R}_+ \Rightarrow 0 + (-\alpha) \in \mathbb{R}_+ \Rightarrow \alpha < 0$. *If* $\alpha < 0 \Rightarrow 0 + (-\alpha) \in \mathbb{R}_+ \Rightarrow -\alpha \in \mathbb{R}_+ \Rightarrow \alpha \in \mathbb{R}_-$

3. *If* $\alpha,\beta \in \mathbb{R}$ *then we have by* 8.15 *the following exclusive possible cases for* $\alpha+(-\beta)$
   a. $\alpha+(-\beta) \in \mathbb{R}_+ \Leftrightarrow \beta < \alpha$
   b. $\alpha+(-\beta) \in \mathbb{R}_- \Leftrightarrow -(\alpha+(-\beta)) \in \mathbb{R}_+ \Leftrightarrow \beta+(-\alpha) \in \mathbb{R}_+ \Leftrightarrow \alpha < \beta$
   c. $\alpha+(-\beta) = 0 \Leftrightarrow \alpha = \beta$
4. *If* $\alpha, \beta \in \mathbb{R}$ *with* $\alpha < \beta$ *then* $\beta+(-\alpha) \in \mathbb{R}_+ \Rightarrow (-\alpha)+(-(-\beta)) \in \mathbb{R}_+ \Rightarrow -\beta < -\alpha$
5. *If* $\alpha, \beta, \gamma \in \mathbb{R}$ *with* $\alpha < \beta$ *then* $\beta+(-\alpha) \in \mathbb{R}_+ \Rightarrow \beta+\gamma+(-\gamma)+(-\alpha) \in \mathbb{R}_+ \Rightarrow (\beta+\gamma)+(-(\alpha+\gamma)) \in \mathbb{R}_+ \Rightarrow \alpha+\gamma < \beta+\gamma$
6. *If* $\alpha, \beta \in \mathbb{R}$ *with* $\alpha < \beta$ *and* $\gamma \in \mathbb{R}_+$ *then* $\beta+(-\alpha) \in \mathbb{R}_+$ *then using* 8.34 *we have* $(\beta+(-\alpha)) \cdot \gamma \in \mathbb{R}_+ \Rightarrow \beta \cdot \gamma+(-\alpha) \cdot \gamma \in \mathbb{R}_+ \Rightarrow (\beta \cdot \gamma)+(-(\alpha \cdot \gamma)) \in \mathbb{R}_+ \Rightarrow \alpha \cdot \gamma < \beta \cdot \gamma$
7. *If* $\alpha, \beta \in \mathbb{R}$ *with* $\alpha < \beta$ *and* $\gamma \in \mathbb{R}_-$ *then* $\beta+(-\alpha) \in \mathbb{R}_+$ *and* $-\gamma \in \mathbb{R}_+$ *then using* 8.34 *we have* $\mathbb{R}_+ \ni (\beta+(-\alpha)) \cdot (-\gamma) = -((\beta+(-\alpha)) \cdot \gamma) = -(\beta \cdot \gamma+(-(\alpha \cdot \gamma))) = (\alpha \cdot \gamma)+(-(\beta \cdot \gamma)) \Rightarrow \beta \cdot \gamma < \alpha \cdot \gamma$
8. *Let* $0 < x$ *consider then the following possible cases for* x^{-1} *are*
   a. $\boldsymbol{(x^{-1} = 0)}$ *then* $1 = x \cdot x^{-1} = 0$ *contradicting* $1 \neq 0$
   b. $\boldsymbol{(x^{-1} < 0)}$ *then using (6)* $1 = x \cdot x^{-1} < 0 \cdot x = 0 \Rightarrow 1 < 0$ *contradicting* $0 < 1$
   c. $\boldsymbol{(0 < x^{-1})}$ *this is the only case left and proves thus (8)*
9. *If* $0 < x < y \underset{(8)}{\Rightarrow} 0 < x^{-1}, y^{-1}$ *and using (6) we have* $1 = x \cdot x^{-1} < y \cdot x^{-1} \Rightarrow 1 < y \cdot x^{-1} \Rightarrow y^{-1} = 1 \cdot y^{-1} < (y \cdot x^{-1}) \cdot y^{-1} = x^{-1} \Rightarrow y^{-1} < x^{-1}$
10. *If* $x \in \mathbb{R}$ *then we have*
    a. $\boldsymbol{(x = 0)} \Rightarrow x^2 = x \cdot x = 0 \cdot 0 = 0 \Rightarrow 0 \leqslant x^2$
    b. $\boldsymbol{(x < 0)} \Rightarrow -0 < -x \Rightarrow 0 < -x \Rightarrow 0 < (-x) \cdot (-x) = (-1)^2 (x \cdot x) = x^2 \Rightarrow 0 < x^2$
    c. $\boldsymbol{(0 < x)} \Rightarrow 0 \cdot x < x \cdot x \Rightarrow 0 < x^2$
11. *If* $x, y \in \mathbb{R}_+ \bigcup \{0\}$ *is such that* $x < y$ *then we have to consider the following cases for* x
    a. $\boldsymbol{(x = 0)}$ *then* $x^2 = 0 \cdot 0 < y \Rightarrow x^2 = 0 \underset{(6)}{=} 0 \cdot y < y \cdot y = y^2 \Rightarrow x^2 < y^2$
    b. $\boldsymbol{(0 < x)}$ *then from* $0 < x < y$ *we have* $x, y \in R_+$ *and using (6) we get* $x \cdot x < y \cdot x$ *and* $x \cdot y < y \cdot y$ *giving* $x \cdot x < y \cdot y \Rightarrow x^2 < y^2$
12. *If* $\alpha \in \mathbb{R}_+$ *then* $0 < \alpha$ *we proceed now by induction on* n *to prove that* $\forall n \in \mathbb{N}_0$ *we have* $0 < \alpha^n$. *So let* $X = \{n \in \mathbb{N}_0 | \text{if } 0 < \alpha \text{ then } 0 < \alpha^n\}$ *then we have:*
    a. $\alpha^0 = 1 > 0 \Rightarrow 0 \in X$
    b. *If* $n \in X$ *then if* $0 < \alpha$ *we have* $0 < \alpha^n$. *Now* $\alpha^{n+1} = \alpha^n \cdot \alpha \underset{0 < \alpha^n, \alpha \in \mathbb{R}_+ \text{and} (6)}{\Rightarrow} 0 < \alpha^n \cdot \alpha = \alpha^{n+1} \Rightarrow n+1 \in X$

    Using mathematical induction (see 4.10*) we have* $X = \mathbb{N}_0$ *proving our assertion.*

13. *Let $\alpha \in \mathbb{R}_+$ with $0 < \alpha < 1$ then if $n \in \mathbb{N}$ we have by (12) already $0 < \alpha^n$. Now to prove $\alpha^n < \alpha$ take $S = \{n < \mathbb{N} | \text{if } \alpha < 1 \text{ then } \alpha^n < \alpha\}$ then we have*

    a. *if $n = 1$ then $\alpha^1 = \alpha < 1$ so that $1 \in S$*

    b. *if $n \in S$ then $\alpha^{n+1} = \alpha^n \cdot \alpha < \alpha$ [as $n \in S \Rightarrow \alpha^n < 1 \underset{0 < \alpha \text{ and } (6)}{\Rightarrow} \alpha^n \cdot \alpha < 1 \cdot \alpha = \alpha$]*

14.

    a. *We prove this by induction so let $\mathcal{S} = \{n \in \mathbb{N} | \alpha < \beta^n\}$ then we have*

        $1 \in \mathcal{S}$. *this follows from $\alpha < \beta = \beta^1$*

        $n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$. *as $n \in \mathcal{S}$ we have $\alpha < \beta^n$ we have by (6) that $\alpha \cdot \beta < \beta \cdot \beta^n = \beta^{n+1}$ and $1 \leq \alpha \Rightarrow \beta \leqslant \alpha \cdot \beta \underset{\alpha < \beta}{\Rightarrow} \alpha < \alpha \cdot \beta < \beta^{n+1}$ proving that $n + 1 \in \mathcal{S}$*

    b. *We prove this by induction so let $\mathcal{S} = \{n \in \mathbb{N} | \alpha \leqslant \beta^n\}$ then we have*

        $1 \in \mathcal{S}$. *this follows from $\alpha \leqslant \beta = \beta^1$*

        $n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$. *as $n \in \mathcal{S}$ we have $\alpha \leqslant \beta^n$ we have by (6) that $\alpha \cdot \beta \leqslant \beta \cdot \beta^n = \beta^{n+1}$ and $1 \leq \alpha \Rightarrow \beta \leqslant \alpha \cdot \beta \underset{\alpha \leqslant \beta}{\Rightarrow} \alpha \leqslant \alpha \cdot \beta \leqslant \beta^{n+1}$ proving that $n + 1 \in \mathcal{S}$*

$\square$

Theorem 8.40. *If $\alpha, \beta \in \mathbb{R}$ and $0 \leqslant \alpha, \beta$ then from $\alpha + \beta = 0$ we have $\alpha = \beta = 0$*

Proof. We have then to consider the following cases

1. $(\mathbf{0} < \boldsymbol{\alpha})$ $\Rightarrow \beta = 0 + \beta < \alpha + \beta = 0 \Rightarrow \beta < 0$ a contradiction
2. $(\mathbf{0} < \boldsymbol{\beta})$ $\Rightarrow \alpha = 0 + \alpha < \alpha + \beta = 0 \Rightarrow \alpha < 0$ a contradiction
3. $(\mathbf{0} = \boldsymbol{\alpha}, \boldsymbol{\beta})$ this is what we have to prove $\square$

Lemma 8.41. *If $r, s \in \mathbb{Q}$ with $r < s$ then $\alpha_r < \alpha_s$ (and thus we have from $r \leqslant s \Rightarrow r = s$ or $r < s \Rightarrow \alpha_r = \alpha_s$ or $\alpha_r < \alpha_s \Rightarrow \alpha_r \leqslant \alpha_s$)*

Proof. If $r, s \in \mathbb{Q}$ with $r < s$ then if $u \in \alpha_r \Rightarrow u < r \underset{r < s}{\Rightarrow} u < s \Rightarrow u \in \alpha_s \Rightarrow \alpha_r \subseteq \alpha_s$. Also $r < s \Rightarrow r \in \alpha_s$ and $r \notin \alpha_r \Rightarrow \alpha_r \neq \alpha_s \Rightarrow \alpha_r < \alpha_s$ $\square$

Theorem 8.42. *The field isomorphism $i_{\mathbb{Q}}: \mathbb{Q} \rightarrow \mathbb{Q}_{\mathbb{R}}$ defined by $q \rightarrow \alpha_q$ is order preserving.*

Proof. This is trivial by using the previous lemma. If $r, s \in \mathbb{Q}$ with $r \leqslant s \Rightarrow \alpha_r \leqslant \alpha_s \Rightarrow i_{\mathbb{Q}}(r) \leqslant i_{\mathbb{Q}}(s)$ $\square$

Theorem 8.43. **($\mathbb{R}$ is conditional complete)** *$\langle \mathbb{R}, \leqslant \rangle$ is conditional complete (see 2.175). Using the definition of conditional completeness this means that $\forall S \subseteq \mathbb{R} \vdash S \neq \emptyset$ for which there exists a $b \in \mathbb{R}$ such that $\forall s \in S$ we have $s \leqslant b$ (existence of a upper bound) we have the existence of $\sup(S)$ (a lowest upper bound). In other words: any nonempty set in $\mathbb{R}$ with a upper bound has a lowest upper bound. Using 2.176 we see that $\forall S \subseteq \mathbb{R} \vdash S \neq \emptyset$ for which there exists a $b \in \mathbb{R}$ such that $\forall s \in S$ we have $b \leqslant s$ (existence of a lower bound) there exists a $\inf(S)$ (in other words every non empty set in $\mathbb{R}$ with a lower bound has greatest lower bound).*

Proof. Let $S \subseteq \mathbb{R}$ with $S \neq \emptyset$ and the existence of a $\beta \in \mathbb{R}$ such that $\forall \alpha \in S$ we have $\alpha \leqslant \beta$. Define then $\gamma = \{r \in \mathbb{Q} | \exists \alpha \in S \vdash r \in \alpha\}$ (in another notation $\gamma = \bigcup_{\alpha \in S} \alpha$). We prove now that γ is a Dedekind's cut and thus that $\gamma \in \mathbb{R}$

1. **($\gamma \neq \emptyset$)** As $S \neq \emptyset$ there exists a $\alpha \in S \subseteq \mathbb{R}$ and by 8.1 , 1 there exists a $r \in \alpha \subseteq \mathbb{Q} \Rightarrow r \in \gamma \Rightarrow \gamma \neq \emptyset$
2. **($\gamma \neq \mathbb{Q}$)** If $r \in \gamma \Rightarrow \exists \alpha \in S$ such that $r \in \alpha$ and as $\alpha \leqslant \beta \Rightarrow \alpha \subseteq \beta \Rightarrow r \in \beta$ and thus we have $\gamma \subseteq \beta$ and as by 8.1 , 2 we have $\beta \neq \mathbb{Q}$ we have also $\gamma \neq \mathbb{Q}$
3. **($\forall r \in \gamma \wedge \forall s \in \mathbb{Q} \backslash \gamma$ we have $r < s$)** So let $r \in \gamma$ and $s \in \mathbb{Q} \backslash \gamma$. From $r \in \gamma$ we have that $\exists \alpha \in S \vdash r \in \alpha$. From $s \in \mathbb{Q} \backslash \gamma$ we have that $\forall \tau \in S$ we have that $s \notin \tau$ so in particular we have $s \notin \alpha \Rightarrow s \in \mathbb{Q} \backslash \alpha$. Using 8.1 , 3 for α we have thus $r < s$.
4. **($\max(\gamma)$ does not exists)** We prove this by contradiction, so assume $m = \max(\gamma)$ exists. Then as $m \in \gamma$ there exists a $\alpha \in S$ such that $m \in \alpha$ and as $\max(\alpha)$ does not exists there exist a $m' \in \alpha$ with $m < m'$, and from $m' \in \alpha \in S$ we derive $m' \in \gamma \Rightarrow m' < m < m' \Rightarrow m' < m'$ a contradiction. So $\max(\gamma)$ does not exists.

Next we prove that γ is a upper bound for S. If $\alpha \in S$ then $\forall r \in \alpha$ we have $r \in \{s \in \mathbb{Q} | \exists \alpha \in S \vdash s \in \alpha\} = \gamma \Rightarrow \alpha \subseteq \gamma \Rightarrow \alpha \leqslant \gamma$. Finally we prove that it is the least upper bound. So if $\gamma' \in \upsilon(S) = \{u \in \mathbb{R} | u \,\text{is}\, a\, \text{upper bound of}\, S\}$ then if $r \in \gamma \Rightarrow \exists \alpha \in S \vdash r \in \alpha$, as $\alpha \leqslant \gamma' \Rightarrow \alpha \subseteq \gamma'$ we have $r \in \gamma'$ and thus we have $\gamma \subseteq \gamma' \Rightarrow \gamma \leqslant \gamma'$ and thus is a least element of $\upsilon(S)$ and thus we have $\gamma = \sup(S)$ □

Theorem 8.44. *Let $S \subseteq \mathbb{R}$ then we have for $-S = \{-s, s \in S\}$ that*

1. *If $\sup(S)$ exists then $\inf(-S)$ exists and $\inf(-S) = -\sup(S)$*
2. *If $\inf(S)$ exists then $\sup(-S)$ exists and $\sup(-S) = -\inf(S)$*

Proof.

1. As $\sup(S)$ exists we have $\forall s \in -S$ that $-s \in S$ so that $-s \leqslant \sup(S)$ or $-\sup(S) \leqslant s$ so $-S$ is bounded below by $-\sup(S)$ proving by the conditional completeness of $\mathbb{R}$ [see 8.43] that
$$\inf(-S)\,\text{exists and}\; -\sup(S) \leqslant \inf(-S). \tag{8.1}$$
Assume now that $-\sup(S) < \inf(-S)$ then $-\inf(-S) < \sup(S)$ so there exist a $s \in S$ such that $-\inf(-S) < s \leqslant \sup(S)$ so that $-s < \inf(-S)$ which as $-s \in -S$ gives $-s < \inf(-S) \leqslant -s$ a contradicton. So we must have that $\inf(-S) \leqslant -\sup(S)$ which together with (8.1) proves that
$$\inf(-S) = -\sup(S)$$
2. As $\inf(S)$ exists we have $\forall s \in -S$ that $-s \in S$ so that $\inf(S) \leqslant -s$ or $s \leqslant -\inf(S)$ so $-S$ is bounded above by $-\inf(S)$ proving by the conditional completeness of $\mathbb{R}$ [see 8.43] that
$$\sup(-S)\,\text{exists and}\; \sup(-S) \leqslant -\inf(S) \tag{8.2}$$

Assume now that $\sup(-S) < -\inf(S)$ then $\inf(S) < -\sup(-S)$ so there exist a $s \in S$ such that $\inf(S) \leqslant s < -\sup(-S)$ so that $\sup(-S) < -s$ which as $-s \in -S$ gives $-s \leqslant \sup(-S) < -s$ a contradicton. So we must have that $-\inf(S) \leqslant \sup(-S)$ which together with (8.2) proves that

$$\sup(-S) = -\inf(S)$$ □

Theorem 8.45. *If $S \subseteq \mathbb{R}$ has a supremum then if $\alpha \in \mathbb{R}$ with $\alpha \geqslant 0$ then $\alpha \cdot S = \{\alpha \cdot s | s \in S\}$ has a supremum where $\sup(\{\alpha \cdot s | s \in S\}) = \alpha \cdot \sup(S)$*

Proof. If $\alpha = 0$ then $\alpha \cdot S = \{0\}$ with $\sup(\alpha \cdot S) = \sup(\{0\}) = 0$ so we only have to prove the case that $\alpha > 0$. By the definition of $\sup(S)$ we have that $\forall s \in S$ that $s \leqslant \sup(S)$ so if $y \in \alpha \cdot S$ then $y = \alpha \cdot s \leqslant \alpha \cdot \sup(S)$ [as $\alpha > 0$] proving that $\alpha \cdot \sup(S)$ is a upper bound for $\alpha \cdot S$. So by the above theorem $\alpha \cdot S$ has a $\sup(\alpha \cdot S)$ with by definition $\sup(\alpha \cdot S) \leqslant \alpha \cdot \sup(S)$. If now $\sup(\alpha \cdot S) < \alpha \cdot \sup(S)$ then $\frac{1}{\alpha} \cdot \sup(\alpha \cdot S) < \sup(S)$ and there exists a $s \in S$ with $\frac{1}{\alpha} \cdot \sup(\alpha \cdot S) < s \leqslant \sup(S)$ giving $\sup(\alpha \cdot S) < \alpha \cdot s \leqslant \alpha \cdot \sup(S)$ which as $\alpha \cdot s \in \alpha \cdot S$ contradicts the fact that $\sup(\alpha \cdot S)$ is the supremum of $\alpha \cdot S$ so we must have $\sup(\alpha \cdot S) = \alpha \cdot \sup(S)$. □

Theorem 8.46. *If $S, T \subseteq \mathbb{R}$ have a supremum then $S + T = \{s + t | s \in S \wedge t \in T\}$ has a supremum with $\sup(S + T) = \sup(S) + \sup(T)$*

Proof. Let $r \in S + T$ then $\exists s \in S, \exists t \in T$ such that $r = s + t$. Using the fact that the supremum is a upper bound we have that $r = s + t \leqslant \sup(S) + \sup(T)$. Using the definition of the supremum we have then that

$$\sup(S + T) \leqslant \sup(S) + \sup(T) \tag{8.3}$$

Assume now that $\sup(S + T) < \sup(S) + \sup(T)$ then if we take $\varepsilon = \sup(S) + \sup(T) - \sup(S + T)$ we have that $0 < \varepsilon$. So $\sup(S) - \frac{\varepsilon}{2} < \sup(S)$ and $\sup(T) - \frac{\varepsilon}{2} < \sup(T)$ and by the definition of the supremum there exists a $s \in S$ and a $t \in S$ such that $\sup(S) - \frac{\varepsilon}{2} < s \leqslant \sup(S) + \frac{\varepsilon}{2}$ and $\sup(T) - \frac{\varepsilon}{2} < t \leqslant \sup(T)$. So $s + t > \sup(S) - \frac{\varepsilon}{2} + \sup(T) - \frac{\varepsilon}{2} = \sup(S) + \sup(T) - \varepsilon = \sup(S) + \sup(T) - \sup(S) - \sup(T) + \sup(S + T) = \sup(S + T)$ or $s + t > \sup(S + T)$ which conflict the fact that the supremum is a upper bound. Hence $\sup(S) + \sup(T) \leqslant \sup(S + T)$ which together with (8.3) proves that

$$\sup(S + T) = \sup(S) + \sup(T)$$ □

Corollary 8.47. *If $S \subseteq \mathbb{R}$ has a supremum and $x \in \mathbb{R}$ then $S + x = \{s + x | s \in S\}$ has a supremum with $\sup(S + x) = \sup(S) + x$*

Proof. Define $T = \{x\}$ then $\sup(T) = x$ exists and $S + T = \{s + t | s \in S \wedge t \in T\} = \{s = x | s \in S\} = S + x$. Hence using 8.46 we have that $\sup(S + x)$ exists and $\sup(S + x) = \sup(S) + \sup(T) = \sup(S) + x$ □

Theorem 8.48. *If $S, T \subseteq \mathbb{R}$ have a infinum then $S + T = \{s + t | s \in S \wedge t \in T\}$ has a infinum with $\inf(S + T) = \inf(S) + \inf(T)$*

Proof. Let $r \in S+T$ then $\exists s \in S, \exists t \in T$ such that $r = s+t$. Using the fact that the infinum is a lower bound we have that $\inf(S) + \inf(T) \leqslant s+t = r$. Using the definition of the infinum we have then that

$$\inf(S) + \inf(T) \leqslant \inf(S+T) \tag{8.4}$$

Assume now that $\inf(S) + \inf(T) < \inf(S+T)$ then if we take $\varepsilon = \inf(S+T) - \inf(S) - \inf(T)$ we have that $0 < \varepsilon$. So $\inf(S) < \inf(S) + \frac{\varepsilon}{2}$ and $\inf(T) < \inf(T) + \frac{\varepsilon}{2}$ and by the definition of the infinum there exists a $s \in S$ and a $t \in S$ such that $\inf(S) \leqslant s < \inf(S) + \frac{\varepsilon}{2}$ and $\inf(T) \leqslant t < \inf(T) + \frac{\varepsilon}{2}$. So $s+t < \inf(S) + \inf(T) + \varepsilon = \inf(S+t) + \inf(T) + \inf(S+T) - \inf(S) - \inf(T) = \inf(S+T)$ or $s+t < \inf(S+T)$ which conflict the fact that the infinum is a lower bound. Hence $\inf(S+T) \leqslant \inf(S) + \inf(T)$ which together with (8.4) proves that

$$\inf(S+T) = \inf(S) + \inf(T)$$

□

Corollary 8.49. *If $S \subseteq \mathbb{R}$ has a infinum and $x \in \mathbb{R}$ then $S + x = \{s + x | s \in S\}$ has a infinum with* $\inf(S+x) = \inf(S) + x$

Proof. Define $T = \{x\}$ then $\inf(T) = x$ exists and $S + T = \{s + t | s \in S \wedge t \in T\} = \{s = x | s \in S\} = S + x$. Hence using 8.48 we have that $\inf(S+x)$ exists and $\inf(S+x) = \inf(S) + \inf(T) = \inf(S) + x$ □

Theorem 8.50. $\langle \mathbb{Q}_{\mathbb{R}}, \leqslant \rangle$ *is not conditional complete*

Proof. Using the fact that $\langle \mathbb{Q}, \leqslant \rangle$ is not conditional complete (see 7.38) we have that there exists a nonempty set $A' \subseteq Q$ which is bounded above by a $u' \in \mathbb{Q}$ such that $\sup(A')$ does not exists. Define now $A = i_{\mathbb{Q}}(A')$ and $u = i_{\mathbb{Q}}(u')$ then we have the following:

1. As A' is not empty there exists a $x \in A'$ then $i_{\mathbb{Q}}(x) \in A \Rightarrow A$ is not empty.

2. If $a \in A \Rightarrow \exists a' \in A'$ such that $a = i_{\mathbb{Q}}(a')$, as u is a upper bound of A' we have that $a' \leqslant u' \underset{i_{\mathbb{Q}} \text{ is order preserving}}{\Rightarrow} a = i_{\mathbb{Q}}(a') \leqslant i_{\mathbb{Q}}(u') = u \Rightarrow A$ is bounded above by u.

So A is a non-empty set that is bounded above. Assume now that $s = \sup(A)$ exists. Take now $s' = i_{\mathbb{Q}}^{-1}(s)$ then we have:

1. If $a' \in A'$ then if $s' < a'$ we have by the order preserving of $i_{\mathbb{Q}}$ and the fact that it is injective that $s = i_{\mathbb{Q}}(s') < i_{\mathbb{Q}}(a') \in A$ so s is not a upper bound of A contradicting the fact that s is the lowest upper bound. So we must have that $a' \leqslant s'$ and as a' was chosen arbitrary s' is a upper bound of A'

2. Let b' be another upper bound of A' and $b = i_{\mathbb{Q}}(b')$. If now $a \in A$ then $a' = i_{\mathbb{Q}}^{-1}(a) \in A'$ [as $i_{\mathbb{Q}}(a') = a \in A$] and thus $a' \leqslant b'$ using then the order preserving of $i_{\mathbb{Q}}$ we have then $a = i_{\mathbb{Q}}(a') \leqslant i_{\mathbb{Q}}(b') = b \Rightarrow b$ is a upper bound of A and as s is the least upper bound we have $s \leqslant b$. If now $b' < s'$ then we have using order preserving and injectivity of $i_{\mathbb{Q}}$ that $b = i_{\mathbb{Q}}(b') < i_{\mathbb{Q}}(s') = s \Rightarrow b < s$ contradicting the fact that s is the lowest upper bound of A. So we must have that $s' \leqslant b'$

From the above we conclude that $s' = \sup(A')$ is a lower upper bound of A' contradicting the fact that we have chosen A' so that no upper bound exists. So we conclude that A has no lowest upper bound and as A is a non empty set that has a upper bound we conclude that $\langle \mathbb{Q}_{\mathbb{R}}, \leqslant \rangle$ is not conditional complete. □

Corollary 8.51. *$\mathbb{Q}_{\mathbb{R}} \subset \mathbb{R}$ so there exists a $r \in \mathbb{R}$ that is not in $\mathbb{Q}$. In other words we have $\mathbb{R} \backslash \mathbb{Q}_{\mathbb{R}} \neq \emptyset$. $\mathbb{R} \backslash \mathbb{Q}_{\mathbb{R}}$ is called the set of irrational numbers.*

Proof. As $\langle \mathbb{Q}_{\mathbb{R}}, \leqslant \rangle$ is not conditional complete (see previous theorem) there exists a non-empty set $A \subseteq \mathbb{Q}_{\mathbb{R}}$ with a upper bound u so that $\sup(A)$ does not exists in $\mathbb{Q}_{\mathbb{R}}$. Because $\mathbb{Q}_{\mathbb{R}} \subseteq \mathbb{R}$ we have $\emptyset \neq A \subseteq \mathbb{R}$ and that A has the upper bound $u \in \mathbb{Q}_{\mathbb{R}} \subseteq \mathbb{R}$. As $\mathbb{R}$ is conditional complete (see 8.43) there exists a lowest upper bound $s = \sup(A)$, now if $s \in \mathbb{Q}_{\mathbb{R}}$ it would be a upper bound of A and if $b \in \mathbb{Q}_{\mathbb{R}}$ of A is another upper bound of A it is also a upper bound of A in $\mathbb{R}$ and thus $s \leqslant b$ so s would be the supremum of A in $\mathbb{Q}_{\mathbb{R}}$ contradicting the fact that $\sup(A)$ does not exists in $\mathbb{Q}_{\mathbb{R}}$. So $s \notin \mathbb{Q}_{\mathbb{R}}$. □

Definition 8.52. $\mathbb{Z}_{\mathbb{R}} = \{\alpha_r | r \in \mathbb{Z}_{\mathbb{Q}}\} \subseteq \mathbb{Q}_{\mathbb{R}} \subseteq \mathbb{R}$. *Note that $0 = \alpha_0, 1 = \alpha_1$ are elements of $\mathbb{Z}_{\mathbb{R}}$.*

Theorem 8.53. *$\langle \mathbb{Z}_{\mathbb{R}}, +, \cdot \rangle$ forms a sub ring of $\langle \mathbb{R}, +, \cdot \rangle$. Also $i_{\mathbb{Q}_{\mathbb{Z}}}: \mathbb{Z} \rightarrow \mathbb{Z}_{\mathbb{R}}$ defined by $x \rightarrow \alpha_{\frac{x}{1}}$ is a ring isomorphism that is also order preserving. So we can consider $\mathbb{Z}_{\mathbb{R}}$ as the set of integer numbers embedded in the set of reals.*

Proof. Let $x, y \in \mathbb{Z}_{\mathbb{R}}$ then we have the existence of a $u, v \in \mathbb{Z}_{\mathbb{Q}} \subseteq \mathbb{Q}$ so that $x = \alpha_u$ and $y = \alpha_v$. Further using the fact that $\mathbb{Z}_{\mathbb{Q}}$ is a sub-ring of $\mathbb{Q}$ (see 7.25) we have that $u + v \in \mathbb{Z}_{\mathbb{Q}}$ and $u \cdot v \in \mathbb{Z}_{\mathbb{Q}}$ and $-u \in \mathbb{Z}_{\mathbb{Q}}$. This gives then

1. $x + y = \alpha_u + \alpha_v \underset{8.27}{=} \alpha_{u+v} \in \mathbb{Z}_{\mathbb{R}}$
2. $x \cdot y = \alpha_u \cdot \alpha_v \underset{8.27}{=} \alpha_u \cdot \alpha_v = \alpha_{u \cdot v} \in \mathbb{Z}_{\mathbb{R}}$
3. $-x = -(\alpha_u) \underset{8.9}{=} \alpha_{-u} \in \mathbb{Z}_{\mathbb{R}}$ (as $-u \in \mathbb{Z}_{\mathbb{Q}}$)
4. $1 \in \mathbb{Z}_{\mathbb{Q}}$ (see definition)
5. $0 \in \mathbb{Z}_{\mathbb{Q}}$ (see definition)

this proves that $\langle \mathbb{Z}_{\mathbb{R}}, +, \cdot \rangle$ is a sub-ring of $\langle \mathbb{R}, +, \cdot \rangle$. Now to prove that $i_{\mathbb{Q}_{\mathbb{Z}}}$ is a order preserving ring isomorphism we first prove that it is a bijection:

1. **(injective)** If $x, y \in \mathbb{Z}$ are such that $\alpha_{\frac{x}{1}} = \alpha_{\frac{y}{1}}$ then by 8.4 we have $\frac{x}{1} = \frac{y}{1} \Rightarrow x \cdot 1 = y \cdot 1 \Rightarrow x = y$ proving injectivity.
2. **(surjective)** If $z \in \mathbb{Z}_{\mathbb{R}}$ then there exists a $r \in \mathbb{Z}_{\mathbb{Q}}$ such that $z = \alpha_r$. As $r \in \mathbb{Z}_{\mathbb{Q}}$ there exists a $u \in \mathbb{Z}$ such that $r = \frac{u}{1}$, but then $z = \alpha_{\frac{u}{1}} = i_{\mathbb{Q}_{\mathbb{Z}}}(u)$ proving surjectivity.

Finally to prove that $i_{\mathbb{Q}_{\mathbb{Z}}}$ is a a order preserving ring isomorphism, note that

1. $\forall x, y \in \mathbb{Z}$ we have that $\frac{x+y}{1} = \frac{x}{1} + \frac{y}{1}$ so $i_{\mathbb{Q}_{\mathbb{Z}}}(x+y) = \alpha_{\frac{x+y}{1}} = \alpha_{\frac{x}{1}+\frac{y}{1}} \underset{8.27}{=} \alpha_{\frac{x}{1}} + \alpha_{\frac{y}{1}} = i_{\mathbb{Q}_{\mathbb{Z}}}(x) + i_{\mathbb{Q}_{\mathbb{Z}}}(y)$

2. $\forall x, y \in \mathbb{Z}$ we have that $\frac{x \cdot y}{1} = \frac{x}{1} \cdot \frac{y}{1}$ so $i_{\mathbb{Q}_{\mathbb{Z}}}(x \cdot y) = \alpha_{\frac{x \cdot y}{1}} = \alpha_{\frac{x}{1} \cdot \frac{y}{1}} \underset{8.27}{=} \alpha_{\frac{x}{1}} \cdot \alpha_{\frac{y}{1}} = i_{\mathbb{Q}_{\mathbb{Z}}}(x) \cdot i_{\mathbb{Q}_{\mathbb{Z}}}(y)$
3. $\forall x, y \in \mathbb{Z}$ with $x \leqslant y$ we have by 7.25 that $\frac{x}{1} = i_{\mathbb{Z}}(x) \leqslant i_{\mathbb{Z}}(y) = \frac{y}{1} \Rightarrow \frac{x}{1} \leqslant \frac{y}{1}$. Using 8.41 we have then $\alpha_{\frac{x}{1}} \leqslant \alpha_{\frac{y}{1}}$ and thus $i_{Q_{\mathbb{Z}}}(x) \leqslant i_{\mathbb{Q}_{\mathbb{Z}}}(y)$.

□

Corollary 8.54. $\mathbb{Z}_{\mathbb{R}}$ *is denumerable*

Proof. As $\mathbb{Z}$ is denumerable (see 6.59) there exists a bijection $\beta: \mathbb{N}_0 \rightarrow \mathbb{Z}$ and as $i_{\mathbb{Q}_{\mathbb{Z}}}: \mathbb{Z} \rightarrow \mathbb{Z}_{\mathbb{R}}$ is a bijection by the previous theorem we have the bijection $i_{\mathbb{Q}_{\mathbb{Z}}} \circ \beta: \mathbb{N}_0 \rightarrow \mathbb{Z}_{\mathbb{R}}$ proving that $\mathbb{Z}_{\mathbb{R}}$ is denumerable. □

Definition 8.55. $\mathbb{N}_{0_{\mathbb{R}}} = \{\alpha_r | r \in \mathbb{N}_{0_{\mathbb{Q}}}\} \subseteq \mathbb{Z}_{\mathbb{R}} \subseteq \mathbb{Q}_{\mathbb{R}} \subseteq \mathbb{R}$ *(for definition of* $\mathbb{N}_{0_{\mathbb{Q}}}$ *see 7.29). Note that* $0 = \alpha_0, 1 = \alpha_1$ *are elements of* $\mathbb{Z}_{\mathbb{R}}$

Theorem 8.56. $\langle \mathbb{N}_{0_{\mathbb{R}}}, + \rangle$ *forms a sub-semi-group of* $\langle \mathbb{R}, + \rangle$, $\langle \mathbb{N}_{0_{\mathbb{R}}}, \cdot \rangle$ *forms a sub-semi-group of* $\langle \mathbb{R}, \cdot \rangle$. *Also* $i_{\mathbb{Q}_{\mathbb{N}_0}}: \mathbb{N}_0 \rightarrow \mathbb{N}_{0_{\mathbb{R}}}$ *defined by* $n \rightarrow \alpha_r$ *where* $r = \frac{i_{\mathbb{N}_0}(n)}{1}$ *is a semi-group isomorphism that is also order preserving (for the addition + and multiplication ·). So we can consider* $\mathbb{N}_{0_{\mathbb{R}}}$ *as the set of natural number embedded in the set of reals. Also* $0 = i_{\mathbb{Q}_{\mathbb{N}_0}}(0)$ *and* $\forall n \in \mathbb{N}_{0_{\mathbb{R}}}$ *we have* $0 \leqslant n$

Proof. Let $x, y \in \mathbb{N}_{0_{\mathbb{R}}}$ then there exists $r, s \in \mathbb{N}_{0_{\mathbb{Q}}}$ such that $x = \alpha_r$ and $y = \alpha_s$. From $r, s \in \mathbb{N}_{0_{\mathbb{Q}}}$ we have the existence of $u, v \in \mathbb{N}_{0_{\mathbb{Z}}}$ such that $r = \frac{u}{1}$ and $s = \frac{v}{1}$. Using the above gives then

1. $x + y = \alpha_r + \alpha_s \underset{8.27}{=} \alpha_{r+s}$. Also $r + s = \frac{u}{1} + \frac{v}{1} = \frac{u \cdot 1 + v \cdot 1}{1 \cdot 1} = \frac{u+v}{1}$ where $u + v \in \mathbb{N}_{0_{\mathbb{Z}}}$ (see 6.22) and this gives that $r + s \in \mathbb{N}_{0_{\mathbb{Q}}} \Rightarrow x + y = \alpha_{r+s} \in \mathbb{N}_{0_{\mathbb{R}}}$
2. $0 \in \mathbb{N}_{0_{\mathbb{R}}}$ (see note in the definition)

we therefore conclude that $\langle \mathbb{N}_{0_{\mathbb{R}}}, + \rangle$ is indeed a sub-semi-group of $\langle \mathbb{R}, + \rangle$.

Next we prove that $\langle \mathbb{N}_{0_{\mathbb{R}}}, \cdot \rangle$ is a sub-semi-group of $\langle \mathbb{R}, \cdot \rangle$

1. $x \cdot y = \alpha_r \cdot \alpha_s \underset{8.27}{=} \alpha_{r \cdot s}$. Also $r \cdot s = \frac{u}{1} \cdot \frac{v}{1} = \frac{u \cdot v}{1}$ where $u \cdot v \in \mathbb{N}_{0_{\mathbb{Z}}}$ (see 6.22) and thus $r \cdot s \in \mathbb{N}_{0_{\mathbb{Q}}} \Rightarrow x \cdot y = \alpha_{r \cdot s} \in \mathbb{N}_{0_{\mathbb{R}}}$
2. $1 \in \mathbb{N}_{0_{\mathbb{R}}}$ (see note in the definition)

Next we prove that $i_{\mathbb{Q}_{\mathbb{N}_0}}$ is a order preserving semi-group isomorphism. First we prove that $i_{\mathbb{Q}_{\mathbb{N}_0}}$ is a bijection:

1. **(Injective)** If n, m are such that $i_{\mathbb{Q}_{\mathbb{N}_0}}(n) = i_{\mathbb{Q}_{\mathbb{N}_0}}(m)$ then if $r = \frac{i_{\mathbb{N}_0}(n)}{1}$, $s = \frac{i_{\mathbb{N}_0}(m)}{1}$ we have that $\alpha_r = \alpha_s \underset{8.4}{\Rightarrow} r = s \Rightarrow i_{\mathbb{N}_0}(n) \cdot 1 = i_{\mathbb{N}_0}(m) \cdot 1 \Rightarrow i_{\mathbb{N}_0}(n) = i_{\mathbb{N}_0}(m) \underset{i_{\mathbb{N}_0} \text{is} \, a \, \text{bijection}}{\Rightarrow} n = m$
2. **(Surjective)** If $x \in \mathbb{N}_{0_{\mathbb{R}}}$ then there exists a $r \in \mathbb{N}_{0_{\mathbb{Q}}}$ such that $x = \alpha_r$, as $r \in \mathbb{N}_{0_{\mathbb{Q}}}$ there exists a $u \in \mathbb{N}_{0_{\mathbb{Z}}}$ such that $r = \frac{u}{1}$, as $u \in \mathbb{N}_{0_{\mathbb{Z}}}$ there exists by 6.22 a $n \in \mathbb{N}_0$ such that $u = i_{\mathbb{N}_0}(n) \Rightarrow r = \frac{i_{\mathbb{N}_0}(n)}{1} \Rightarrow x = \alpha_r = i_{\mathbb{Q}_{\mathbb{N}_0}}(n)$

Next we prove that it is a semi-group isomorphism

1. If $n,m\in \mathbb{N}_0$ then $n+m\in \mathbb{N}_0$ and by 6.22 we have $i_{\mathbb{N}_0}(n+m)=i_{\mathbb{N}_0}(n)+i_{\mathbb{N}_0}(m)$ and thus $\frac{i_{\mathbb{N}_0}(n+m)}{1}=\frac{i_{\mathbb{N}_0}(n)+i_{\mathbb{N}_0}(m)}{1}=\frac{i_{\mathbb{N}_0}(n)}{1}+\frac{i_{\mathbb{N}_0}(m)}{1}$. So we have that $i_{\mathbb{Q}_{\mathbb{N}_0}}(n+m)=\alpha_{\frac{i_{\mathbb{N}_0}(n+m)}{1}}=\alpha_{\frac{i_{\mathbb{N}_0}(n)}{1}+\frac{i_{\mathbb{N}_0}(m)}{1}}\underset{8.27}{=}\alpha_{\frac{i_{\mathbb{N}_0}(n)}{1}}+\alpha_{\frac{i_{\mathbb{N}_0}(m)}{1}}=i_{\mathbb{Q}_{\mathbb{N}_0}}(n)+i_{\mathbb{Q}_{\mathbb{N}_0}}(m)$
2. If $n,m\in \mathbb{N}_0$ then $n\cdot m\in \mathbb{N}_0$ and by 6.22 we have $i_{\mathbb{N}_0}(n\cdot m)=i_{\mathbb{N}_0}(n)\cdot i_{\mathbb{N}_0}(m)$ and thus $\frac{i_{\mathbb{N}_0}(n\cdot m)}{1}=\frac{i_{\mathbb{N}_0}(n)\cdot i_{\mathbb{N}_0}(m)}{1}=\frac{i_{\mathbb{N}_0}(n)}{1}\cdot\frac{i_{\mathbb{N}_0}(m)}{1}$. So we have that $i_{\mathbb{Q}_{\mathbb{N}_0}}(n\cdot m)=\alpha_{\frac{i_{\mathbb{N}_0}(n\cdot m)}{1}}=\alpha_{\frac{i_{\mathbb{N}_0}(n)}{1}\cdot\frac{i_{\mathbb{N}_0}(m)}{1}}\underset{8.27}{=}\alpha_{\frac{i_{\mathbb{N}_0}(n)}{1}}\cdot\alpha_{\frac{i_{\mathbb{N}_0}(m)}{1}}=i_{\mathbb{Q}_{\mathbb{N}_0}}(n)\cdot i_{\mathbb{Q}_{\mathbb{N}_0}}(m)$

Next to prove order preserving:

1. If $n,m\in\mathbb{N}_0\vdash n\leqslant m$ then using 6.31 we have $i_{\mathbb{N}_0}(n)\leqslant i_{\mathbb{N}_0}(m)$ and then using 7.25 we have $\frac{i_{\mathbb{N}_0}(n)}{1}=i_{\mathbb{Z}}(i_{\mathbb{N}_0}(n))\leqslant i_{\mathbb{Z}}(i_{\mathbb{N}_0}(m))=\frac{i_{\mathbb{N}_0}(m)}{1}$. So finally using 8.41 we have $i_{\mathbb{Q}_{\mathbb{N}_0}}(n)=\alpha_{\frac{i_{\mathbb{N}_0}(n)}{1}}\leqslant\alpha_{\frac{i_{\mathbb{N}_0}(m)}{1}}=i_{\mathbb{Q}_{\mathbb{N}_0}}(m)$ proving that $i_{\mathbb{Q}_{\mathbb{N}_0}}$ is order preserving.

Next if $0\in\mathbb{N}_0$ then $i_{\mathbb{N}_0}(0)=\sim[(s(0),1)]=\sim[(1,1)]=0\in\mathbb{Z}\Rightarrow\frac{i_{\mathbb{N}_0}(0)}{1}=\frac{0}{1}=0\in\mathbb{Q}\Rightarrow i_{\mathbb{Q}_{\mathbb{N}_0}}(0)=\alpha_0=0\in\mathbb{R}$.

Also if $1\in\mathbb{N}_0\Rightarrow i_{\mathbb{N}_0}(1)=\sim[(s(1),1)]=\sim[(2,1)]=1\in\mathbb{Z}\Rightarrow\frac{i_{\mathbb{N}_0}(1)}{1}=\frac{1}{1}=1\in\mathbb{Q}\Rightarrow i_{\mathbb{Q}_{\mathbb{N}_0}}(1)=\alpha_1=1\in\mathbb{R}$.

Finally to prove that $\forall n\in\mathbb{N}_{0_{\mathbb{R}}}$ we have $0\leqslant n$. If $n\in\mathbb{N}_{0_{\mathbb{R}}}$ then by the above there exists a $m\in\mathbb{N}_0$ such that $i_{\mathbb{Q}_{\mathbb{N}_0}}(m)=n$. Using 4.47 we have $0\leqslant m$ and using the order preserving properties of $i_{\mathbb{Q}_{\mathbb{N}_0}}$ we have $0=i_{\mathbb{Q}_{\mathbb{N}_0}}(0)\leqslant i_{\mathbb{Q}_{\mathbb{N}_0}}(m)=n\Rightarrow 0\leqslant n$ $\square$

As $u=i_{\mathbb{Q}_N}$ is a bijection.

Corollary 8.57. $\mathbb{N}_{0_{\mathbb{R}}}$ *is denumerable*

Theorem 8.58. *We have the following properties concerning* $\mathbb{N}_{0_{\mathbb{R}}}$ *and* $\mathbb{Z}_{\mathbb{R}}$ *(here* $-\mathbb{N}_{0_{\mathbb{R}}}=\{n\vdash -n\in\mathbb{N}_{0_{\mathbb{R}}}\}\subseteq\mathbb{Z}_{\mathbb{R}}$*)*

1. $\mathbb{Z}_{\mathbb{R}}=\mathbb{N}_{0_{\mathbb{R}}}\bigcup(-\mathbb{N}_{0_{\mathbb{R}}})$ *and* $\mathbb{N}_{0_{\mathbb{R}}}\bigcap(-\mathbb{N}_{0_{\mathbb{R}}})=\{0\}$
2. *If* $n\in\mathbb{Z}_{\mathbb{R}}$ *and* $0\leqslant n\Rightarrow n\in\mathbb{N}_{0_{\mathbb{R}}}$ *(so* $\mathbb{N}_{0_{\mathbb{R}}}$ *is indeed the set of positive integers)*
3. $\langle\mathbb{N}_{0_{\mathbb{R}}},+\rangle$ *is a semi-group of* $\langle\mathbb{Z}_{\mathbb{R}},+\rangle$
4. $\langle\mathbb{N}_{0_{\mathbb{R}}},\cdot\rangle$ *is a sub-semi-group of* $\langle\mathbb{Z}_{\mathbb{R}},\cdot\rangle$

Proof.

1. As $\mathbb{N}_{0_{\mathbb{R}}},-\mathbb{N}_{0_{\mathbb{R}}}\subseteq\mathbb{Z}_{\mathbb{R}}$ we have $\mathbb{N}_{0_{\mathbb{R}}}\bigcup(-\mathbb{N}_{0_{\mathbb{R}}})\subseteq\mathbb{Z}_{\mathbb{R}}$. If $x\in\mathbb{Z}_{\mathbb{R}}$ then $x=\frac{n}{1}$ where $n\in\mathbb{Z}$ then using 6.24 we have either
   a. $n\in\mathbb{N}_{0_{\mathbb{Z}}}\Rightarrow x\in\mathbb{N}_{0_{\mathbb{R}}}$
   b. $n\in-\mathbb{N}_{0_{\mathbb{Z}}}\Rightarrow -n\in\mathbb{N}_{0_{\mathbb{Z}}}\Rightarrow -x=\frac{-n}{1}\in\mathbb{N}_{0_{\mathbb{R}}}\Rightarrow x\in-\mathbb{N}_{0_{\mathbb{R}}}$

   and thus $x\in\mathbb{N}_{0_{\mathbb{Z}}}\bigcup(-\mathbb{N}_{0_{\mathbb{Z}}})\Rightarrow\mathbb{Z}_{\mathbb{R}}\subseteq\mathbb{N}_{0_{\mathbb{R}}}\bigcup(-\mathbb{N}_{0_{\mathbb{R}}})\Rightarrow\mathbb{Z}_{\mathbb{R}}=\mathbb{N}_{0_{\mathbb{R}}}\bigcup(-\mathbb{N}_{0_{\mathbb{R}}})$. Finally if $x\in\mathbb{N}_{0_{\mathbb{R}}}\bigcap(-\mathbb{N}_{0_{\mathbb{R}}})$ then $0\leqslant x$ (as $x\in\mathbb{N}_{0_{\mathbb{R}}}$) also $-x\in\mathbb{N}_{0_{\mathbb{R}}}\Rightarrow -x=\frac{n}{1}$ where $n\in\mathbb{N}_{0_{\mathbb{Z}}}\underset{6.28}{\Rightarrow}0\leqslant n\underset{0\leqslant 1\Rightarrow 0\leqslant n\cdot 1}{\Rightarrow}\operatorname{sign}(-x)=1\Rightarrow 0\leqslant -x\Rightarrow x\leqslant -0=0\Rightarrow 0\leqslant x\leqslant 0\Rightarrow x=0\Rightarrow\mathbb{N}_{0_{\mathbb{R}}}\bigcap(-\mathbb{N}_{0_{\mathbb{R}}})=\{0\}$

2. If $n \in \mathbb{Z}_{\mathbb{R}}$ and $0 \leqslant n$ then $\exists q \in \mathbb{Z}_{\mathbb{Q}}$ such that $n = \alpha_q$ and thus $\exists z \in \mathbb{Z}$ such that $q = \frac{z}{1}$. As $0 = \alpha_0 \leqslant \alpha_q$ we must have $0 \leqslant q$ [if $q < 0$ then by 8.41 we have $\alpha_q < \alpha_0$ contradicting $\alpha_0 \leqslant \alpha_q$]. As $0 \leqslant q \Rightarrow \operatorname{sign}(q) = 1 \Rightarrow 0 \leqslant z \cdot 1 = z \underset{6.28}{\Rightarrow} z \in \mathbb{N}_{0_{\mathbb{Z}}} \Rightarrow q = \frac{z}{1} \in \mathbb{N}_{0_{\mathbb{Q}}} \Rightarrow n = \alpha_q \in \mathbb{N}_{0_{\mathbb{R}}} \Rightarrow n \in \mathbb{N}_{0_{\mathbb{R}}}$
3. We can use the same prove as in 8.56
4. We can use the same prove as in 8.56

□

Theorem 8.59. *$\langle \mathbb{N}_{0_{\mathbb{R}}} \leqslant \rangle$ is well-ordered*

Proof. If $B \subseteq \mathbb{N}_{0_{\mathbb{R}}}$ is a nonempty subset of $\mathbb{N}_{0_{\mathbb{R}}}$ then $\exists b \in B \underset{i_{\mathbb{Q}_{\mathbb{N}_0}} \text{is surjective}}{\Rightarrow} \exists n \in \mathbb{N}_0 \vdash b = i_{\mathbb{Q}_{\mathbb{N}_0}}(n) \Rightarrow i_{\mathbb{Q}_{\mathbb{N}_0}}^{-1}(B)$ is a nonempty subset of $\mathbb{N}_0$. As $\langle \mathbb{N}_0, \leqslant \rangle$ is well-ordered (see 4.52) there exist a least element $m = \min\left(i_{\mathbb{Q}_{\mathbb{N}_0}}^{-1}(B)\right)$ so $m \in i_{\mathbb{Q}_{\mathbb{N}_0}}^{-1}(B)$ and $\forall n \in i_{\mathbb{Q}_{\mathbb{N}_0}}^{-1}(B)$ we have $m \leqslant n$. Now $M = i_{\mathbb{Q}_{\mathbb{N}_0}}(m) \in B$ (as $m \in i_{\mathbb{Q}_{\mathbb{N}_0}}^{-1}(B)$) and if $N \in B \underset{i_{\mathbb{Q}_{\mathbb{N}_0}} \text{is surjective}}{\Rightarrow} \exists n \in \mathbb{N}_0 \vdash i_{\mathbb{Q}_{\mathbb{N}_0}}(n) = N \in B \Rightarrow n \in i_{\mathbb{Q}_{\mathbb{N}_0}}^{-1}(B) \Rightarrow m \leqslant n \underset{i_{\mathbb{Q}_{\mathbb{N}_0}} \text{is order preserving}}{\Rightarrow} M = i_{\mathbb{Q}_{\mathbb{N}_0}}(m) \leqslant i_{\mathbb{Q}_{\mathbb{N}_0}}(n) = N \Rightarrow M \leqslant N$ or $M = \min(B)$ and B has a least element. □

Lemma 8.60. *If $n \in \mathbb{N}_{0_{\mathbb{R}}}$ and $0 < n \Rightarrow 1 \leqslant n$.*

Proof. As $n \in \mathbb{N}_{0_{\mathbb{R}}}$ there exists by surjectivity of $i_{\mathbb{Q}_{\mathbb{N}_0}}$ a $n' \in \mathbb{N}_0$ such that $i_{\mathbb{Q}_{\mathbb{N}_0}}(n') = n$. Using 4.47 we have $0 \leqslant n'$, if $n' = 0 \Rightarrow n = i_{\mathbb{Q}_{\mathbb{N}_0}}(n') = i_{\mathbb{Q}_{\mathbb{N}_0}}(0) = 0$ contradicting $0 < n$, this leaves us to conclude that $0 < n'$. Using 4.51 we have then $1 = s(n) \leqslant n'$ and using the order preserving properties of $i_{\mathbb{Q}_{\mathbb{N}_0}}$ and the fact that $i_{\mathbb{Q}_{\mathbb{N}_0}}(1) = 1$ (see 8.56 that $1 = i_{\mathbb{Q}_{\mathbb{N}_0}}(1) \leqslant i_{\mathbb{Q}_{\mathbb{N}_0}}(n') = n \Rightarrow 1 \leqslant n$ □

Theorem 8.61. (Archimedean Property) *If $x, y \in \mathbb{R}$ with $0 < x$ then $\exists n \in \mathbb{N}_{0_{\mathbb{R}}}$ such that $y < n \cdot x$*

Proof. We have the following cases to consider for y

1. **($\boldsymbol{y \leqslant 0}$)** then take $1 \in \mathbb{N}_{0_{\mathbb{R}}}$ then we have $y \leqslant 0 < x \Rightarrow y < x = 1 \cdot x$
2. **($\mathbf{0 < y}$)** we prove the theorem here by contradiction, so assume that $\forall n \in \mathbb{N}_{0_{\mathbb{R}}}$ we have $n \cdot x \leqslant y$. Define $A = \{n \cdot x | n \in \mathbb{N}_{0_{\mathbb{R}}}\}$ then $\forall t \in A$ we have $t \leqslant y$ and thus y is a upper bound of A. From conditional completeness (see 9.43) $\sup(A)$ exists. As $x > 0$ we have using 9.41 that $-x < 0 \Rightarrow \sup(A) - x < \sup(A)$ as $\sup(A)$ is the least upper bound of A $\sup(A) - x$ is not a upper bound of A so there exists a $n \in \mathbb{N}_{0_{\mathbb{R}}}$ such that $\sup(A) - x < n \cdot x \Rightarrow \sup(A) < n \cdot x + x = (n+1) \cdot x$, now as $n, 1 \in \mathbb{N}_{0_{\mathbb{R}}}$ we have using 8.56) that $n + 1 \in \mathbb{N}_{0_{\mathbb{R}}} \Rightarrow (n+1) \cdot x \in A$ and from $\sup(A) < (n+1) \cdot x \in A$ we have that $\sup(A)$ is not a upper bound of A and reach thus a contradiction and the initial hypothesis is false. So we must have that $\exists n \in \mathbb{N}_{0_{\mathbb{R}}}$ such that $y < n \cdot x$. □

The following theorem shows some of the consequences of the Archimedean property of $\mathbb{R}$

Theorem 8.62. *Let $x \in \mathbb{R}$ then the following holds:*

1. $\exists n \in \mathbb{N}_{0_{\mathbb{R}}} \vdash x < n$
2. $\exists n \in \mathbb{Z}_{\mathbb{R}} \vdash n \leqslant x < n + 1$
3. *If* $0 < x$ *then* $\exists n \in \mathbb{N}_{0_{\mathbb{R}}} \backslash \{0\}$ *such that* $\frac{1}{n} < x$
4. *If* $x \geqslant 0$ *then* $\exists n \in \mathbb{N}_{0_{\mathbb{R}}} \backslash \{0\}$ *such that* $n - 1 \leqslant x < n$

Proof.

1. *As* $0 < 1 \in \mathbb{R}$ *we have by the Archimedean property (see 8.61) that* $\exists n \in \mathbb{N}_{0_{\mathbb{R}}}$ *such that* $x < 1 \cdot n = n$
2. *We have for* y *two possibilities*

   $\boldsymbol{0 \leqslant x}$**.** *then by (1) we have* $A = \{n \in \mathbb{N}_0 | y < n\} \neq \emptyset$ *and by the well ordering of* $\mathbb{N}_{0_{\mathbb{R}}}$ *there exists a least element* $m \in A \Rightarrow x < n$ *then as* $m - 1 < m$ *we have* $\neg(x < m - 1) \Rightarrow m - 1 \leqslant x$*. So if we take* $n = m - 1 \in \mathbb{N}_{0_{\mathbb{R}}} \subseteq \mathbb{Z}_{\mathbb{R}}$ *we have* $n \leqslant x < n + 1$

   $\boldsymbol{x < 0}$**.** *then we have* $0 < -x$ *then we have that* $A = \{n \in \mathbb{N}_0 | -y \leqslant n\} \neq \emptyset$ *and there exists a least element* $m \in A \Rightarrow -x \leqslant m$ *and as* $m - 1 < m$ *we have* $m - 1 < -x \Rightarrow -m \leqslant x < -m + 1$*, take* $n = -m \in \mathbb{Z}_{\mathbb{R}}$ *then* $n \leqslant x < n + 1$
3. *Using the Archimedean property on* 1 *(see 8.61) we have* $\exists n \in \mathbb{N}_{0_{\mathbb{R}}}$ *such that* $1 < x \cdot n$*, as* $n = 0$ *would give* $1 = 0$ *a contradiction we have that* $0 < n$ *and using 8.39 we have then* $0 < \frac{1}{n} \Rightarrow \frac{1}{n} < x$
4. *Consider* $A = \{n \in \mathbb{N}_{0_{\mathbb{R}}} \backslash \{0\} | y < n\} \subseteq \mathbb{N}_{0_{\mathbb{R}}} \backslash \{0\} \subseteq \mathbb{N}_{0_{\mathbb{R}}}$*, using (1) there exists a* $n \in \mathbb{N}_{0_{\mathbb{R}}}$ *such that* $0 \leqslant x < n \Rightarrow n \in A \Rightarrow A \neq \emptyset$*. Using the fact* $\mathbb{N}_{0_{\mathbb{R}}}$ *is well-ordered (see 8.59) we have that* $m = \min(A)$ *exist. Then as* $m - 1 < m$ *we have that* $m - 1 \notin A$ *we have now the following possibilities*

   a. $\boldsymbol{(m - 1 \in \mathbb{N}_{0_{\mathbb{R}}} \backslash \{0\})}$ *as* $m - 1 \notin A$ *we must have* $m - 1 \leqslant x < m$ *(as* $m \in A$*)*

   b. $\boldsymbol{(m - 1 \notin \mathbb{N}_{0_{\mathbb{R}}} \backslash \{0\})}$ *as* $m \in A \Rightarrow m \neq 0$ *we have using 8.56 that* $0 < m \underset{8.60}{\Rightarrow} 1 \leqslant m$ *and as* $1, m \in \mathbb{N}_{0_{\mathbb{R}}} \subseteq \mathbb{Z}_{\mathbb{R}}$ *we have by 8.53 that* $0 = 1 - 1 \leqslant m - 1 \in \mathbb{Z}_{\mathbb{R}} \underset{8.58}{\Rightarrow} m - 1 \in \mathbb{N}_{0_{\mathbb{R}}} \underset{m-1 \notin \mathbb{N}_{0_{\mathbb{R}}} \backslash \{0\}}{\Rightarrow} m - 1 = 0$ *and as* $m \in A, 0 \leqslant x$ *we have* $m - 1 = 0 \leqslant x < m \Rightarrow m - 1 \leqslant x < m$ □

Corollary 8.63. *Let* $x, y \in \mathbb{R}$ *then we have*

1. *If* $\forall \varepsilon \in \mathbb{R}_+$ *we have* $x \leqslant y + \varepsilon$ *then* $x \leqslant y$
2. *If* $\forall \varepsilon \in \mathbb{R}_+$ *we have* $x < y + \varepsilon$ *then* $x \leqslant y$
3. *If* $\forall n \in \mathbb{N}_{0_{\mathbb{R}}} \backslash \{0\}$ *we have* $x \leqslant y + \frac{1}{n}$ *then* $x \leqslant y$
4. *If* $\forall n \in \mathbb{N}_{0_{\mathbb{R}}} \backslash \{0\}$ *we have* $x < y + \frac{1}{n}$ *then* $x \leqslant y$
5. *If* $\forall \varepsilon \in \mathbb{R}_+, a \in \mathbb{R} \vdash 0 \leqslant a$ *we have* $x \leqslant y + \varepsilon \cdot a$ *then* $x \leqslant y$
6. *If* $\forall \varepsilon \in \mathbb{R}_+, a \in \mathbb{R} \vdash 0 \leqslant a$ *we have* $x < y + \varepsilon \cdot a$ *then* $x \leqslant y$

7. *If* $\forall n\in\mathbb{N}_{0_\mathbb{R}}\backslash\{0\}$, $a\in\mathbb{R}\vdash 0\leqslant a$ *we have* $x\leqslant y+\frac{1}{n}\cdot a$ *then* $x\leqslant y$
8. *If* $\forall n\in\mathbb{N}_{0_\mathbb{R}}\backslash\{0\}$, $a\in\mathbb{R}\vdash 0\leqslant a$ *we have* $x<y+\frac{1}{n}\cdot a$ *then* $x\leqslant y$

Proof.

1. Assume that $y<x\Rightarrow 0<x-y$ then by the above theorem (3) there exists a $n\in\mathbb{N}_{0_\mathbb{R}}\backslash\{0\}$ such that for $\varepsilon=\frac{1}{n}$ we have $\varepsilon<x-y\Rightarrow y+\varepsilon<x\leqslant y+\varepsilon\Rightarrow\varepsilon<\varepsilon$ a contradiction so we must have $x\leqslant y$
2. $x<y+\varepsilon\Rightarrow x\leqslant y+\varepsilon\underset{(1)}{\Rightarrow}x\leqslant y$
3. If $\varepsilon\in\mathbb{N}_{0_\mathbb{R}}\backslash\{0\}$ then by the previous theorem there exists a $0<\frac{1}{n}<\varepsilon\Rightarrow$ $x\leqslant y+\frac{1}{n}<y+\varepsilon\underset{(2)}{\Rightarrow}x\leqslant y$
4. $x<y+\frac{1}{n}\Rightarrow x\leqslant y+\frac{1}{n}\underset{(4)}{\Rightarrow}x\leqslant y$
5. Given $\varepsilon\in\mathbb{R}_+$ take $\varepsilon'=\frac{\varepsilon}{a+1}>0$ then from $x\leqslant y+\varepsilon'\cdot a=y+\frac{a}{a+1}\cdot\varepsilon<$ $y+\varepsilon\underset{(2)}{\Rightarrow}x\leqslant y$
6. $x<y+\varepsilon\cdot a\Rightarrow x\leqslant y+\varepsilon\cdot a\underset{(5)}{\Rightarrow}x\leqslant y$
7. Given $n\in\mathbb{N}_{0_\mathbb{R}}\backslash(0)$ then as $0<\frac{1}{n\cdot(a+1)}$ there exists by the above theorem a $m\in\mathbb{N}_{0_\mathbb{R}}\backslash\{0\}$ such that $\frac{1}{m}<\frac{1}{n\cdot(a+1)}\Rightarrow\frac{1}{m}\cdot a<\frac{1}{n}\cdot\frac{a}{a+1}<\frac{1}{n}$ so that $x\leqslant$ $y+\frac{1}{m}\cdot a<y+\frac{1}{n}\underset{(3)}{\Rightarrow}x\leqslant y$
8. $x<y+\frac{1}{n}\cdot a\Rightarrow x\leqslant y+\frac{1}{n}\cdot a\underset{(7)}{\Rightarrow}x\leqslant y$ □

Theorem 8.64. (Density Theorem) *If* $x,y\in\mathbb{R}$ *such that* $x<y$ *then* $\exists r\in\mathbb{Q}_\mathbb{R}$ *and* $\exists i\in\mathbb{R}\backslash\mathbb{Q}_\mathbb{R}$ *such that* $x<r<y$ *and* $x<i<y$

Proof. First we prove that $\exists r\in\mathbb{Q}$ such that $x<r<y$, we have then the following possible cases for x

1. **(0<x)** then from $x<y$ we have $0<y-x\Rightarrow(y-x)^{-1}$ is defined and by 9.41 we have $0<(y-x)^{-1}$, then using the Archimedean property (see 9.54) and $0<1$ we have $\exists n\in\mathbb{N}_{0_\mathbb{R}}$ such that $0<(y-x)^{-1}<n\cdot 1=n\underset{0<(y-x)^{-1}}{\Rightarrow}1<n\cdot(y-x)\Rightarrow$ $1<n\cdot y-n\cdot x\Rightarrow n\cdot x+1<n\cdot y$ and $0<n$. As $0<x$ and $n\in\mathbb{N}_{0_\mathbb{R}}\Rightarrow 0\leqslant n\Rightarrow$ $0\leqslant n\cdot x\underset{9.55}{\Rightarrow}\exists m\in\mathbb{N}_{0_\mathbb{R}}\backslash\{0\}$ such that $m-1\leqslant n\cdot x<m\Rightarrow m\leqslant n\cdot x+1<n\cdot y$. Using the fact that $n,m\in\mathbb{N}_{0_\mathbb{R}}\subseteq\mathbb{Z}_\mathbb{R}\subseteq\mathbb{Q}_\mathbb{R}$ so as $\mathbb{Q}_\mathbb{R}$ is a sub-field in $\mathbb{R}$ (see 8.28) and $0<n$ we have $r=m\cdot n^{-1}\in\mathbb{Q}_\mathbb{R}$, from $0<n$ we have $0<n^{-1}$ and thus $r=m\cdot n^{-1}<n\cdot y\cdot n^{-1}=y\Rightarrow r<y$. Also from $n\cdot x<m$ we have $x=n^{-1}\cdot n\cdot x<n^{-1}\cdot m=r\Rightarrow x<r$. So we have $x<r<y$.
2. **(0=x)** then $0<y$ and by 9.55 there exists a $n\in\mathbb{N}_{0_\mathbb{R}}\backslash\{0\}\subseteq\mathbb{Q}_\mathbb{R}\Rightarrow 0<$ $n\underset{9.41}{\Rightarrow}0<n^{-1}$ such that for $n^{-1}\in\mathbb{Q}_\mathbb{R}$ (as $\mathbb{Q}_\mathbb{R}$ is a sub-field see 8.28) we have $0<n^{-1}<x$. So taking $r=n^{-1}$ we have $0=x<r<y$
3. **(x<0)** We have now the following possibilities for y
   a. $(\mathbf{0<y})$ then if we take $r=0\in\mathbb{Q}_\mathbb{R}\Rightarrow x<r<y$

b. $(\boldsymbol{y \leqslant 0})$ then $x < y \leqslant 0 \Rightarrow 0 \leqslant -y < -x$ and we can use (1) or (2) to find a $r' \in \mathbb{Q}_{\mathbb{R}}$ such that $-y < r' < -x \Rightarrow x < -r' < y$ so using $r = -r' \in \mathbb{Q}_{\mathbb{R}}$ we have $x < r < y$

Next we prove the existence of a $i \in \mathbb{R} \backslash \mathbb{Q}_{\mathbb{R}}$ such that $x < i < y$. As $\mathbb{R} \backslash \mathbb{Q}_{\mathbb{R}} \neq \emptyset$ (see 8.51) there exists a $z \in \mathbb{R} \backslash \mathbb{Q}_{\mathbb{R}}$ then from $x < y$ we have $x+z < y+z$ and by (1),(2) and (3) we have that there exists a $r \in \mathbb{Q}_{\mathbb{R}}$ such that $x+z < r < y+z \Rightarrow x < r-z < y$. So if $i = r-z$ then we have $x < i < y$. Now if $i \in \mathbb{Q}_{\mathbb{R}}$ then as $\mathbb{Q}_{\mathbb{R}}$ is a sub-field (see 8.28) and $r \in \mathbb{Q}_{\mathbb{R}}$ we have $z = -i + r \in \mathbb{Q}_{\mathbb{R}} \in \mathbb{R} \backslash \mathbb{Q}_{\mathbb{R}}$. □

Definition 8.65. *If $x \in \mathbb{R}$ then $|x|$ is defined by $|x| = \begin{cases} x \text{ if } 0 \leqslant x \\ -x \text{ if } x < 0 \end{cases}$ and is called the absolute value*

Theorem 8.66. *The absolute value $||$ satisfies the following*

1. *$\forall x \in \mathbb{R}$ we have $|x| \geqslant 0$*
2. *$\forall x, \alpha \in \mathbb{R}$ we have $|\alpha \cdot x| = |\alpha| \cdot |x|$*
3. *$\forall x, y \in \mathbb{R}$ we have $|x + y| \leqslant |x| + |y|$*
4. *$\forall x \in \mathbb{R}$ we have $|x| = 0 \Leftrightarrow x = 0$*

Proof.

1. *This is trivial based on the definition.*
2. *Given $x, \alpha \in \mathbb{R}$ we have the following cases:*
   a. $(\boldsymbol{0 \leqslant x, \alpha})$ *then* $0 \leqslant x \cdot \alpha$ *and* $|\alpha \cdot x| = \alpha \cdot x = |\alpha| \cdot |x|$
   b. $(\boldsymbol{0 \leqslant x, \alpha < 0})$ *then* $\alpha \cdot x < 0$ *and* $|\alpha \cdot x| = -(\alpha \cdot x) = (-\alpha) \cdot x = |\alpha| \cdot |x|$
   c. $(\boldsymbol{x < 0, 0 \leqslant \alpha})$ *then* $\alpha \cdot x < 0$ *and* $|\alpha \cdot x| = -(\alpha \cdot x) = \alpha \cdot (-x) = |\alpha| \cdot |x|$
   d. $(\boldsymbol{x < 0, \alpha < 0})$ *then* $0 < (-\alpha) \cdot (-x) = \alpha \cdot x$ *and* $|\alpha \cdot x| = \alpha \cdot x = (-\alpha) \cdot (-x) = |\alpha| \cdot |x|$
3. *Given $x, y \in \mathbb{R}$ we have*
   a. $(\boldsymbol{0 \leqslant x, y})$ *then* $0 \leqslant x + y$ *and* $|x + y| = x + y = |x| + |y| \Rightarrow |x + y| \leqslant |x| + |y|$
   b. $(\boldsymbol{0 \leqslant x, y < 0})$ *then we have the following cases*
      i. $(\boldsymbol{0 \leqslant x + y})$ *then* $|x + y| = x + y \underset{y<0 \Rightarrow y<0<-y=|y|}{\Rightarrow} x + y < x + |y| = |x| + |y| \Rightarrow |x + y| \leqslant |x| + |y|$
      ii. $(\boldsymbol{x + y < 0})$ *then* $|x + y| = -x - y \underset{0 \leqslant x \Rightarrow -x \leqslant 0 \leqslant x}{\Rightarrow} -x - y \leqslant x - y \underset{-y=|y|}{=} x + |y| = |x| + |y| \Rightarrow |x + y| \leqslant |x| + |y|$
   c. $(\boldsymbol{x < 0, 0 \leqslant y})$ *then if we take* $x' = y$ *and* $y' = x$ *we have* $|x + y| = |y + x| = |x' + y'| \leqslant |x'| + |y'|$ *(case (b)) and thus* $|x + y| \leqslant |x| + |y|$
   d. $(\boldsymbol{x < 0, y < 0})$ *then* $x + y < 0 \Rightarrow |x + y| = -(x + y) = (-x) + (-y) = |x| + |y| \Rightarrow |x + y| \leqslant |x| + |y|$
4. *If* $x = 0 \underset{0 \leqslant x}{\Rightarrow} |x| = x = 0$, *if* $|x| = 0$ *then if* $x < 0$ *we have* $0 < -x = |x| = 0$ *a contradiction so we must have* $0 \leqslant x$ *but then* $x = |x| = 0 \Rightarrow x = 0$ □

We prove now some useful theorems needed for the calculations of limits, first some definitions to simplify our notation.

Definition 8.67. *Let* $n \in \mathbb{N}_0$ *and* $x \in \mathbb{R}$ *then* $n \cdot x = \begin{cases} 0 \text{ if } n=0 \\ x + (n-1) \cdot x \text{ if } n > 1 \end{cases}$. *Hence* $0.x = 0$ *(last* 0 *is the neutral element in* $\mathbb{R}$*),* $1 \cdot x = x + 0 \cdot x = x + 0 = x$*,* $2 \cdot x = x + 1 \cdot x = x + x$*, ...*

The above operation is consistent with the multiplication in $\mathbb{R}$ as the following theorem proves.

Theorem 8.68. *Let* $n \in \mathbb{N}_0$ *then* $n \cdot x = i_{\mathbb{Q}_{\mathbb{N}_0}}(n) \cdot x$ *hence if* 1 *is the multiplicative unit in* $\mathbb{R}$ *we have* $n \cdot 1 = i_{\mathbb{Q}_{\mathbb{N}_0}}(n)$

Proof. By 8.56 we have that $i_{\mathbb{Q}_{\mathbb{N}_0}}(0) = 0 \in \mathbb{R}$ and $i_{\mathbb{Q}_{\mathbb{N}_0}}(1) = 1$. We prove now the theorem by induction so let $\mathcal{S} = \{n \in \mathbb{N}_0 | n \cdot x = i_{\mathbb{Q}_{\mathbb{N}_0}}(n) \cdot x\}$ then we have

$0 \in \mathcal{S}$. As $0 \cdot x = 0 = i_{\mathbb{Q}_{\mathbb{N}_0}}(0) = i_{\mathbb{Q}_{\mathbb{N}_0}}(0) \cdot x$

$n \in \mathcal{S} = < n+1 \in \mathcal{S}$. we have that $(n+1) \cdot x \underset{n+1>0}{=} x + ((n+1)-1) \cdot x \underset{n \in \mathcal{S}}{=} x + i_{\mathbb{Q}_{\mathbb{N}_0}}(n) \cdot x = i_{Q_{\mathbb{N}_0}}(1) \cdot x + i_{\mathbb{Q}_{\mathbb{N}_0}}(n) \cdot x = i_{\mathbb{Q}_{\mathbb{N}_0}}(n+1) \cdot x = x$ $\square$

Motivated by the above theorem and the fact that we already use the sames symbols for the neutral element and multiplicative unit in $\mathbb{R}$ and $\mathbb{N}_0$ we introduce the following notation.

Notation 8.69. *If* $n \in \mathbb{N}_0$ *then we note* $i_{\mathbb{Q}_{\mathbb{N}_0}}(n)$ *also by* n*. So if in a operation (sum, multiplication or comparison) we mix elements of* $\mathbb{N}_0$ *and elements of* $\mathbb{R}$ *we implicitly assume that we have applied* $i_{\mathbb{Q}_{\mathbb{N}_0}}$ *to the elements of* $\mathbb{N}_0$*.*

Theorem 8.70. $\forall n \in \mathbb{N}_0$ *we have that* $n < 2^n$ *[and thus by 8.62 we have also* $\forall a \in \mathbb{R}$ $\exists n \in \mathbb{N}_0$ *such that* $a < n < 2^n$*]*

Proof. This is proved by induction so let $S = \{n \in \mathbb{N}_0 | n < 2^n\}$. Then we have that

1. If $n = 0$ then $0 < 1 = 2^0$ so that $0 \in S$
2. If $n \in S$ then $i_{\mathbb{Q}_{\mathbb{N}_0}}(n) < 2^n$ then consider the following cases for $n+1$ (as $0 < n+1$)
   a. **$(n+1=1)$** then $n+1 = 1 < 2 = 2^1 = 2^{n+1} \Rightarrow n+1 \in S$
   b. **$(1 < n+1)$** then $2 \leqslant n+1 \Rightarrow 1 \leqslant n$ so that $n+1 \leqslant n+n = 2 \cdot n < 2 \cdot 2^n = 2^{n+1}$ proving that $n+1 \in S$ $\square$

Theorem 8.71. *Let* $x \in \mathbb{R}$ *with* $x > 1$ *then* $x^n - 1 \geqslant n \cdot (x-1)$ $\forall n \in \mathbb{N}_0$

Proof. We prove this by induction, so let $S = \{n \in \mathbb{N}_0 | x^n - 1 \geqslant n \cdot (x-1)\}$ then we have

1. if $n = 0$ then $x^n - 1 = x^0 - 1 = 1 - 1 = 0 \geqslant 0 = 0 \cdot (x-1) = n \cdot (x-1) \Rightarrow 0 \in S$
2. If $n \in S$ then $x^{n+1} - 1 = x \cdot x^{n+1} - 1 = x \cdot (x^n - 1) + (x-1) \geqslant x \cdot (n \cdot (x-1)) + (x-1) = (x-1) \cdot (x \cdot n + 1) >_{x>1 \Rightarrow n \cdot x \geqslant n} (x-1) \cdot (n+1) \Rightarrow n+1 \in S$ $\square$

Theorem 8.72. *If $N \in \mathbb{N}_{0_{\mathbb{R}}} \setminus \{0\}$ and $x \in \mathbb{R}$ with $x > 1$ then there exists a $n \in \mathbb{N}_{0_{\mathbb{R}}} \setminus \{0\}$ such that $x^n > N$*

Proof. Consider the following cases for $N \in \mathbb{N}_{0_{\mathbb{R}}} \setminus \{0\}$

1. **($N = 1$)** then $N = 1 < x = x^1$ proving the theorem
2. **($N > 1$)** take $\delta = \frac{N-1}{x-1} > 0$ then by the Archimedean property (8.62) there exists a $n \in \mathbb{N}_{0_{\mathbb{R}}}$ such that $\delta < n$ as by definition of δ we have $N - 1 = \delta \cdot (x-1) \leqslant n \cdot (x-1) < x^n - 1$ (using the previous theorem) so that $N < x^n$ □

Theorem 8.73. *If $x \in \mathbb{R}$ with $0 < x < 1$ and $n, m \in \mathbb{N}_{0_{\mathbb{R}}}$ with $n < m$ then $x^m < x^n$*

Proof. We prove this by induction, so let $m \in \mathbb{N}_{0_{\mathbb{R}}}$ and take $S_m = \{n \in \mathbb{N} | x^{m+n} < x^m\}$ then we have

1. if $n = 1$ then $x^{m+n} = x^{n+1} = x \cdot (x^n) <_{x < 1 \Rightarrow x \cdot x^n < x^n} x^n \Rightarrow 1 \in S_m$
2. if $n \in S_m$ then $x^{m+(n+1)} = x \cdot x^{m+n} <_{x<1} x^{m+n} <_{n \in S_m} x^m$

Using mathematical induction we have then that $S_M = \mathbb{N}$. If now $n, m \in \mathbb{N}_{0_{\mathbb{R}}}$ with $n < m$ then $k = m - n > 0$ or $k \in \mathbb{N} = S_n$ so that $x^m = x^{n+k} < x^n$ □

Theorem 8.74. *If $\varepsilon \in \mathbb{R}_+$ and $x \in \mathbb{R}$ such that $0 < x < 1$ then $\exists N \in \mathbb{N}_{0_{\mathbb{R}}} \setminus \{0\}$ such that $0 < x^n < \varepsilon \quad$ if $n \geqslant N$*

Proof. As $\varepsilon > 0$ we have that $\frac{1}{\varepsilon}$ is defined and by the Archimedean property (see 8.62) there exists a $n \in \mathbb{N}_{0_{\mathbb{R}}}$ such that $0 < \frac{1}{\varepsilon} < n$, as $0 < x < 1$ we have that $1 < \frac{1}{x}$ so that by 8.72 there exists a $N \in \mathbb{N}_{0_{\mathbb{R}}} \setminus \{0\}$ such that $n < x^N$ and thus we have that $0 < \frac{1}{\varepsilon} < \left(\frac{1}{x}\right)^N = \frac{1}{x^N} \underset{x, \varepsilon > 0}{\Rightarrow} 0 < x^N < \varepsilon$. If now we have $n \geqslant N$ then by the previous theorem we have $0 < x^n \leqslant x^N < \varepsilon$ □

8.4 Square root

Theorem 8.75. *The function $^2\colon \{x \in \mathbb{R} | 0 \leqslant x\} \to \{x \in \mathbb{R} | 0 \leqslant x\}$ defined by $x \to x^2$ (where $0 \leqslant x^2$ by 9.41 so this is indeed a function) is a bijection.*

Proof.

1. **(injectivity)** If $x, y \geqslant 0$ is such that $x^2 = y^2$ then either
   a. **($x = 0$)** then $0 = x^2 = y^2 \Rightarrow y^2 = 0 \underset{9.41}{\Rightarrow} y = 0 \Rightarrow x = y$
   b. **($y = 0$)** then $0 = y^2 = x^2 \Rightarrow x = 0 \Rightarrow x = y$
   c. **($0 < x, y$)** then if we assume that $x \neq y$ then we have the following cases
      i. **($x < y$)** then $x^2 < y \cdot x \underset{x < y \Rightarrow y \cdot x < y^2}{\Rightarrow} x^2 < y^2$ contradicting $x^2 = y^2$
      ii. **($y < x$)** then similar we have $y^2 < x^2$ contradicting $x^2 = y^2$

   In all the cases we have a contradiction so we must have $x = y$.

2. **(surjectivity)** Given $y \in \{x \in \mathbb{R} | 0 \leqslant x\}$ then $0 \leqslant y$ and we have the following cases :

   a. $(\boldsymbol{y = 0})$ then $y = 0 = 0 \cdot 0 = 0^2 \Rightarrow 0^2 = y$

   b. $(\boldsymbol{y = 1})$ then $1^2 = 1 \cdot 1 = 1 = y \Rightarrow 1^2 = y$

   c. $(\boldsymbol{0 < y \neq 1})$ take then $S_y = \{t \in \mathbb{R} | 0 \leqslant t \wedge t^2 \leqslant y\}$. Then as $0^2 = 0 < y$ we have $0 \in S_y$ and thus $0 \notin S_y$. Now as $y \neq 1$ we have the following possibilities for y

      i. $(\boldsymbol{y < 1})$ Take now $t \in S_y$ and suppose $1 < t$. Then from $1 < t$ and $0 < 1 < t \Rightarrow 0 < t$ and 9.41 we have $t \leqslant t^2$. From $y < 1$ and $1 < t$ we have then $y < t$ and using $t \leqslant t^2$ we have then $y < t^2$ which contradicts $t \in S_y \Rightarrow t^2 \leqslant y$. So we must have $t \leqslant 1$ and thus we have that 1 is a upper bound of S_t.

      ii. $(\boldsymbol{1 < y})$ Take now $t \in S_y$ and suppose $y < t$. Then from $1 < y$ and $y < t$ we have $1 < t$. Using $1 < t \Rightarrow 0 < t$ and 9.41 on $1 < t$ we have $1 \cdot t < t \cdot t \Rightarrow t < t^2 \underset{y<t}{\Rightarrow} y < t^2$ contradicting $t \in S_y \Rightarrow t^2 \leqslant y$. So we must have $t \leqslant y$ and thus we have that y is a upper bound of S_y.

   Using (i) and (ii) above we see that S_y has a upper bound and is not empty. From the conditional completeness on the reals (see 9.43) we have that there exists a lowest upper bound $s_y = \sup(S_y)$. For S_y we consider now again the two possible cases for $y \neq 1$

      i. $(\boldsymbol{1 < y})$ then $1^2 = 1 < y \Rightarrow 1 \in S_y$ and thus S_y contains a t with $0 < t$

      ii. $(\boldsymbol{y < 1})$ then from $0 < y$ and 9.41 we have $y \cdot y < 1 \cdot y \Rightarrow y^2 < y \Rightarrow y \in S_y$ and thus S_y contains a t with $0 < t$.

   From the above (i) and (ii) we have ,as $\forall t \in S_y$ that $t \leqslant s_y = \sup(S_y)$, that

$$0 < s_y. \tag{8.5}$$

   Now from $0 < s_y$ there exists by 9.55 a $\varepsilon_0 \in \mathbb{R}$ such that $0 < \varepsilon_0 < s_y$. Now for all $\varepsilon \in \mathbb{R}$ with $0 < \varepsilon < s_y$ we have $0 < s_y - \varepsilon < s_y < s_y + \varepsilon$, then using 9.41 we have $(s_y - \varepsilon)^2 < s + y \cdot (s_y - \varepsilon) \wedge (s_y - \varepsilon) \cdot s_y < s_y^2$ and $s_y^2 < s_y \cdot (s_y + \varepsilon) \wedge s_y \cdot (s_y + \varepsilon) < (s_y + \varepsilon)^2$ giving finally $(s_y - \varepsilon)^2 < s_y^2 < (s_y + \varepsilon)^2$. As s_y is a upper bound of S_y and $s_y < s_y + \varepsilon$ we must have $s_y + \varepsilon \notin S_y$ and as $0 < s_y < s_y + \varepsilon$ we must have that $y < (s_y + \varepsilon)^2$. Using the fact that s_y is the lowest bound on S_y and the fact that $s_y - \varepsilon < s_y$ there exists a $f \in S_y$ such that $s_y - \varepsilon < f \leqslant s_\varepsilon$. Using 9.41 we have then that $f \cdot (s_y - \varepsilon) < f^2 \wedge (s_y - \varepsilon)^2 < f \cdot (s_y - \varepsilon) \Rightarrow (s_y - \varepsilon)^2 < f^2$ and as $f \in S_y$ we have $f^2 \leqslant y$ giving us $(s_y - \varepsilon)^2 < f^2 \leqslant y < (s_y + \varepsilon)^2$. To summarize we have proved that if $0 < \varepsilon_0 < s_y$ that

$$(s_y - \varepsilon)^2 < s_y^2 < (s_y + \varepsilon)^2 \tag{8.6}$$

$$(s_y - \varepsilon)^2 < y < (s_y + \varepsilon)^2 \tag{8.7}$$

Using 9.41 on 8.7 gives

$$-(s_y+\varepsilon)^2 < -y < -(s_y-\varepsilon)^2 \tag{8.8}$$

and adding 8.6 and 8.8 gives

$$(s_y-\varepsilon)^2-(s_y+\varepsilon)^2 < s_y^2 - y < (s_y+\varepsilon)^2-(s_y-\varepsilon)^2 \tag{8.9}$$

As $(s_y+\varepsilon)^2-(s_y-\varepsilon)^2 = s_y^2+2\cdot s_y\cdot\varepsilon+\varepsilon^2-s_y^2+2\cdot s_y\cdot\varepsilon-\varepsilon^2 = 4\cdot\varepsilon\cdot s_y$ and using 8.9 we have then

$$-4\cdot\varepsilon\cdot s_y < s_y^2-y < 4\cdot\varepsilon\cdot s_y \text{ or } -4\cdot\varepsilon\cdot s_y < y-s_y^2 < 4\cdot\varepsilon\cdot s_y \tag{8.10}$$

For s_y^2-y we have now the following possible cases to consider

i. $\boldsymbol{(s_y^2-y<0)}$ Take then $\delta=y-s_y^2$ then $0<\delta$. Take now $\varepsilon=\min\left(\frac{\delta}{4\cdot s_y},\varepsilon_0\right)$ (which is well defined as $0<s_y$ see 8.5) then we have $0<\varepsilon\leqslant\varepsilon_0<s_y$ so we have by 8.10 that $\delta=y-s_y^2<4\cdot\varepsilon\cdot s_y$. As by definition of ε we have $\varepsilon\leqslant\frac{\delta}{4\cdot s_y}\Rightarrow 4\cdot\varepsilon\cdot s_y\leqslant\delta<4\cdot\varepsilon\cdot s_y$ and we reach a contradiction.

ii. $\boldsymbol{(0<s_y^2-y)}$ Take then $\delta=s_y^2-y$ then we have (see 8.10) $0<\delta$. Take now $\varepsilon=\min\left(\frac{\delta}{4\cdot s_y},\varepsilon_0\right)$ then we have $0<\varepsilon\leqslant\varepsilon_0<s_y$, so we have by 8.10 that $\delta=s_y^2-y<4\cdot\varepsilon\cdot s_y$. As by definition of ε we have $\varepsilon\leqslant\frac{\delta}{4\cdot s_y}\Rightarrow 4\cdot\varepsilon\cdot s_y\leqslant\delta<4\cdot\varepsilon\cdot s_y$ and we reach a contradiction.

iii. $\boldsymbol{(s_y^2-y=0)}$ Then $s_y^2=y$

So (iii) is the only possible case [as (i) and (ii) leads to a contradiction] we must have $s_y^2=y$.

So in cases (a), (b) and (c) we have found a t such that $t^2=y$ proving surjectivity. □

Definition 8.76. *Using the above theorem we have that the function* $^2\colon\{x\in\mathbb{R}|0\leqslant x\}\to\{x\in\mathbb{R}|0\leqslant x\}$ *has a inverse mapping. This mapping is called the square root function and is noted by* $\sqrt{\ }$. *So* $\sqrt{\ }\colon\{x\in\mathbb{R}|0\leqslant x\}\to\{x\in\mathbb{R}|0\leqslant x\}$ *is mapping such that* $\sqrt{\ }\circ{}^2=i_{\{x\in\mathbb{R}|0\leqslant x\}}={}^2\circ\sqrt{\ }$ *or* $\forall x\in\{x\in\mathbb{R}|0\leqslant x\}$ *we have* $(\sqrt{x})^2=x$ *and* $\sqrt{x^2}=x$

Note 8.77. The requirement that $0\leqslant x$ to have $\sqrt{x^2}=x=\sqrt{x}^2$ is really necessary as $^2\colon\mathbb{R}\to\{x\in\mathbb{R}|0\leqslant x\}$ is definitely not a injection and thus a bijection as $(-1)^2=1=1^2$

Theorem 8.78. $\sqrt{\ }\colon\{x\in\mathbb{R}|0\leqslant x\}\to\{x\in\mathbb{R}|0\leqslant x\}$ *is a strictly increasing function*

Proof. If $x,y\in\{x\in\mathbb{R}|0\leqslant x\}$ is such that $x<y$. Assume then that $\sqrt{y}\leqslant\sqrt{x}$ then by 8.39 we have $(\sqrt{y})^2\leqslant(\sqrt{x})^2\Rightarrow y\leqslant x$ contradicting $x<y$ so we must have $\sqrt{x}<\sqrt{y}$ □

Theorem 8.79. *If* $x, y \in \{x \in \mathbb{R} | 0 \leqslant x\}$ *then* $\sqrt{x \cdot y} = \sqrt{x} \cdot \sqrt{y}$

Proof. As $(\sqrt{x \cdot y})^2 = x \cdot y = (\sqrt{x})^2 \cdot (\sqrt{y})^2 = (\sqrt{x} \cdot \sqrt{y})^2$ we have by the fact that ^{2}: $\{x \in \mathbb{R} | 0 \leqslant x\} \rightarrow \{x \in \mathbb{R} | 0 \leqslant x\}$ is a bijection and thus injective so we have $\sqrt{x \cdot y} = \sqrt{x} \cdot \sqrt{y}$ □

Theorem 8.80. *Given* $x, y \in \{x \in \mathbb{R} | 0 \leqslant x\}$ *then we have* $\sqrt{x + y} \leqslant \sqrt{x} + \sqrt{y}$

Proof. We prove this by contradiction, so assume that $\sqrt{x} + \sqrt{y} < \sqrt{x + y} \underset{8.39}{\Rightarrow} -\sqrt{x + y} < -(\sqrt{x} + \sqrt{y})$ and by multiplying by $\sqrt{x} + \sqrt{y} \geqslant 0$ we have $-\sqrt{x + y} \cdot (\sqrt{x} + \sqrt{y}) < -(\sqrt{x} + \sqrt{y}) \cdot (\sqrt{x} + \sqrt{y}) = -(x + y + 2 \cdot \sqrt{x} \cdot \sqrt{y})$. Now we have by 8.39 that $0 \leqslant (\sqrt{x + y} - (\sqrt{x} + \sqrt{y}))^2 = (\sqrt{x + y})^2 + (\sqrt{x} + \sqrt{y})^2 - 2 \cdot \sqrt{x + y} \cdot (\sqrt{x} + \sqrt{y}) = x + y + (\sqrt{x})^2 + (\sqrt{y})^2 + 2 \cdot \sqrt{x} \cdot \sqrt{y} - 2 \cdot \sqrt{x + y} \cdot (\sqrt{x} + \sqrt{y}) = 2 \cdot x + 2 \cdot y + 2 \cdot \sqrt{x} \cdot \sqrt{y} - 2 \cdot \sqrt{x + y} \cdot (\sqrt{x} + \sqrt{y}) \underset{\text{multiply by } 0 < \frac{1}{2}}{\Rightarrow} 0 \leqslant x + y + \sqrt{x} \cdot \sqrt{y} - \sqrt{x + y} \cdot (\sqrt{x} + \sqrt{y}) < x + y + \sqrt{x} \cdot \sqrt{y} - (x + y + 2 \cdot \sqrt{x}\,\sqrt{y}) = -\sqrt{x} \cdot \sqrt{y} \leqslant 0 \Rightarrow 0 < 0$ a contradiction. So we must have $\sqrt{x + y} \leqslant \sqrt{x} + \sqrt{y}$ □

Chapter 9
The complex numbers

9.1 Definition and arithmetic's

Definition 9.1. *The space $\langle \mathbb{C}, +, \cdot \rangle$ of complex numbers is defined by: (note that $+, \cdot$ has different meanings in $\mathbb{C}$ and $\mathbb{R}$)*

1. $\mathbb{C} = \mathbb{R} \times \mathbb{R}$
2. $+: \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ *by* $((x, y), (x', y')) \rightarrow (x, y) + (x', y') \underset{\text{defined}}{=} (x + x', y + y')$
3. $\cdot: \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ *by* $((x, y), (x', y')) \rightarrow (x, y) \cdot (x', y') \underset{\text{defined}}{=} (x \cdot x' - y \cdot y', x \cdot y' + x' \cdot y)$

Theorem 9.2. *$\langle \mathbb{C}, +, . \rangle$ forms a field. The unit element is $(1, 0)$ noted also as 1, the neutral element is $(0, 0)$ noted as 0,$-(x, y) = (-x, -y)$ and if $(x, y) \neq (0, 0)$ then $(x, y)^{-1} = \left(\frac{x^2 \cdot y}{x^3 \cdot y + x \cdot y^3}, \frac{-y^2 \cdot x}{x^3 \cdot y + x \cdot y^3} \right)$ if $x \neq 0 \wedge y \neq 0$, $(x, y)^{-1} = (x^{-1}, 0)$ if $x \neq 0$, $y = 0$ and $(x, y)^{-1} = (0, -y^{-1})$ if $x = 0, y \neq 0$*

Proof. First we prove that $\langle \mathbb{C}, + \rangle$ forms a abelian group using the fact that $\langle \mathbb{R}, +, \cdot \rangle$ is a field.

1. **(associativity)** $\forall (x, y), (x', y'), (x'', y'') \in \mathbb{C}$ we have $(x, y) + ((x', y') + (x'', y'')) = (x, y) + (x' + x'', y' + y'') = (x + (x' + x''), y + (y' + y'')) = ((x + x') + x'', (y + y') + y'') = (x + x', y + y') + (x'', y'') = ((x, y) + (x', y')) + (x'', y'')$
2. **(neutral element)** $\forall (x, y) \in \mathbb{C}$ we have $(x, y) + (0, 0) = (x + 0, y + 0) = (x, y) = (x + 0, y + 0) = (x, y) + (0, 0)$
3. **(inverse element)** $\forall (x, y) \in \mathbb{C}$ we have $(x, y) + (-x, -y) = (x + (-x), y + (-y)) = (0, 0) = (-x + x, -y + y) = (-x, -y) + (x, y)$. So $(-x, -y)$ is the additive inverse of (x, y).
4. **(commutativity)** $\forall (x, y), (x', y') \in \mathbb{C}$ we have $(x, y) + (x', y') = (x + x', y + y') = (x' + x, y' + y) = (x', y') + (x, y)$

Next we prove the rest of the axioms of a field

1. **(distributive)** $\forall (x, y), (x', y'), (x'', y'')$ we have $(x, y) \cdot ((x', y') + (x'', y'')) = (x, y) \cdot (x' + x'', y' + y'') = (x \cdot (x' + x'') - y \cdot (y' + y''), x \cdot (y' + y'') + y \cdot (x' + x'')) = (x \cdot x' + x \cdot x'' - y \cdot y' - y \cdot y'', x \cdot y' + x \cdot y'' + y \cdot x' + y \cdot x'') = (x \cdot x' - y \cdot y' + x \cdot x'' - y \cdot y'', x \cdot y' + y \cdot x' + x \cdot y'' + y \cdot x'') = (x \cdot x' - y \cdot y', x \cdot y' + y \cdot x') + (x \cdot x'' - y \cdot y'', x \cdot y'' + y \cdot x'') = (x, y) \cdot (x', y') + (x, y) \cdot (x'', y'')$

2. **(neutral element)** $\forall (x,y)\in\mathbb{C}$ we have $(1,0)\cdot(x,y)=(1\cdot x-0\cdot y,1\cdot y+0\cdot x)=(x,y)$ and $(x,y)\cdot(1,0)=(x\cdot 1+y\cdot 0,x\cdot 0+y\cdot 1)=(x,y)$
3. **(commutative)** $\forall (x,y),(x',y')\in\mathbb{C}$ we have $(x,y)\cdot(x',y')=(x\cdot x'-y\cdot y', x\cdot y'+y\cdot x')=(x'\cdot x-y'\cdot y,y'\cdot x+x'\cdot y)=(x'\cdot x-y'\cdot y,x'\cdot y+y'\cdot x)=(x', y')\cdot(x,y)$
4. **(associative)** $\forall (x,y),(x',y'),(x'',y'')$ then we have $(x,y)\cdot((x',y')\cdot(x'', y''))=(x,y)\cdot(x'\cdot x''-y'\cdot y'',x'\cdot y''+y'\cdot x'')=(x\cdot(x'\cdot x''-y'\cdot y'')-y\cdot(x'\cdot y''+y'\cdot x''),x\cdot(x'\cdot y''+y'\cdot x'')+y\cdot(x'\cdot x''-y'\cdot y''))=(x\cdot x'\cdot x''-x\cdot y'\cdot y''-y\cdot x'\cdot y''-y\cdot y'\cdot x'',x\cdot x'\cdot y''+x\cdot y'\cdot x''+y\cdot x'\cdot x''-y\cdot y'\cdot y'')=((x\cdot x'-y\cdot y')\cdot x''-(x\cdot y'+y\cdot x')\cdot y'',(x\cdot x'-y\cdot y')\cdot y''+(x\cdot y'+y\cdot x')\cdot x'')=(x\cdot x'-y\cdot y',x\cdot y'+y\cdot x')\cdot(x'',y'')=((x,y)\cdot(x',y'))\cdot(x'',y'')$
5. **(multiplicative inverse)** Given $(x,y)\in\mathbb{C}$ with $(x,y)\neq(0,0$, then we have 3 cases to consider

   $\boldsymbol{x\neq 0\wedge y\neq 0.}$ Now $x^3\cdot y+x\cdot y^3=x\cdot(x^2\cdot y+y^3)=x\cdot y\cdot(x^2+y^2)$ now using 9.41 we have $0<x^2,y^2\Rightarrow 0<(x^2+y^2)\underset{x,y\neq o}{\Rightarrow} x\cdot y\cdot(x^2+y^2)\Rightarrow x^3\cdot y+x\cdot y^3\neq 0$. So $\left(\frac{x^2\cdot y}{x^3\cdot y+x\cdot y^3},\frac{-y^2\cdot x}{x^3\cdot y+x\cdot y^3}\right)$ is defined and $(x, y)\cdot\left(\frac{x^2\cdot y}{x^3\cdot y+x\cdot y^3},\frac{-y^2\cdot x}{x^3\cdot y+x\cdot y^3}\right)=\left(\frac{x\cdot x^2\cdot y+y\cdot y^2\cdot x}{x^3\cdot y+x\cdot y^3},\frac{x\cdot -y^2\cdot x+y\cdot x^2\cdot y}{x^3\cdot y+x\cdot y^3}\right)=\left(\frac{x^3\cdot y+x\cdot y^3}{x^3\cdot y+x\cdot y^3},\frac{-x^2\cdot y^2+x^2\cdot y^2}{x^3\cdot y+x\cdot y^3}\right)=(1,0)\underset{(3)}{=}\left(\frac{x^2\cdot y}{x^3\cdot y+x\cdot y^3},\frac{-y^2\cdot x}{x^3\cdot y+x\cdot y^3}\right)\cdot(x, y)$

   $\boldsymbol{x\neq 0\wedge y=0.}$ then $(x,y)\cdot(x^{-1},0)=(x,0)\cdot(x^{-1},0)=(x\cdot x^{-1}-0\cdot 0, x\cdot 0+0\cdot x^{-1})=(x\cdot x^{-1},0)=1$

   $\boldsymbol{x=0\wedge y\neq 0.}$ then $(x,y)\cdot(0,y^{-1})=(0,y)\cdot(0,y^{-1})=(0\cdot 0-y\cdot(-y^{-1}), 0\cdot y^{-1}+y\cdot 0)=(1,0)$ □

Notation 9.3. *If $x,y\in\mathbb{C}$ with $y\neq 0$ then we note $x\cdot y^{-1}$ as $\frac{x}{y}$ hence as $x^{-1}=1\cdot x^{-1}=\frac{1}{y}$ we have also that $y^{=1}=\frac{1}{y}$*

Definition 9.4. *Given $(x,y)\in\mathbb{C}$ we define $\overline{(x,y)}=(x,-y)$*

Lemma 9.5. *If $z,z'\in\mathbb{C}$ then $\overline{z+z'}=\bar{z}+\bar{z}'$, $\overline{z\cdot z'}=\bar{z}\cdot\bar{z}'$ and $\bar{\bar{z}}=z$*

Proof. If $z=(x,y),z'=(x',y')$ then

1. $\overline{(x,y)+(x',y')}=\overline{(x+x',y+y')}=(x+x',-y-y')=(x,-y)+(x',-y')=\overline{(x,y)}+\overline{(x',y')}$
2. $\overline{(x,y)\cdot(x',y')}=\overline{(x\cdot x'-y\cdot y',x\cdot y'+y\cdot x')}=(x\cdot x'-y\cdot y',-x\cdot y'-y\cdot x')=(x\cdot x'-(-y)\cdot(-y'),x\cdot(-y')+(-y)\cdot x')=(x,-y)\cdot(x',-y')=\overline{(x,y)}\cdot\overline{(x',y')}$
3. $\bar{\bar{z}}=\overline{\overline{(x,y)}}=\overline{(x,-y)}=(x,y)=z$ □

Definition 9.6. *As $\langle\mathbb{C},\cdot\rangle$ is a abelian semi-group we have by 4.22 that given a $a\in\mathbb{C}$ and $n\in\mathbb{N}_0$ that there exists a a^n such that*

$$\begin{aligned} a^0 &= 1\\ a^{n+1} &= a^n\cdot a\underset{\text{abelian}}{=} a\cdot a^n\end{aligned}$$

Theorem 9.7. *If $n, n' \in \mathbb{N}_0$ and $a \in \mathbb{C}$ then $a^{n'+n} = a^{n'} \cdot a^n$*

Proof. We prove this by induction on n. So let $X = \{n \in \mathbb{N}_0 | a^{n'+n} = a^{n'} \cdot a^n\}$ then we have

1. If $n = 0$ then $a^{n'+n} = a^{n'+0} = a^{n'} = a^{n'} \cdot 1 = a^{n'} \cdot a^0 \Rightarrow 0 \in X$
2. If $n \in X$ then $a^{n'+(n+1)} = a^{(n'+n)+1} = a^{(n'+n)} \cdot a \underset{n \in X}{=} (a^{n'} \cdot a^n) \cdot a = a^{n'} \cdot (a^n \cdot a) = a^{n'} \cdot a^{n+1}$ and thus $n+1 \in X$

Using mathematical induction (see 4.10) we have $X = \mathbb{N}_0$ proving the theorem □

Theorem 9.8. *If $n \in \mathbb{N}_0$ and $a, b \in \mathbb{C}$ then we have $(a \cdot b)^n = a^n \cdot b^n$*

Proof. We prove this by induction so take $\mathcal{S} = \{n \in \mathbb{N}_0 | (a \cdot b)^n = a^n \cdot b^n\}$ then we have

$n = 0$. then $(a \cdot b)^0 = 1 = 1 \cdot 1 = a^0 \cdot b^0$

$n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$. then $(a \cdot b)^{n+1} = (a \cdot b)^n \cdot (a \cdot b) = (a^n \cdot b^n) \cdot (a^n \cdot a) \cdot (b^n \cdot b) = a^{n+1} \cdot b^{n+1}$ □

Corollary 9.9. *If $z \in \mathbb{C}$ then $\forall n \in \mathbb{N}$ we have that $\overline{z^n} = \bar{z}^n$*

Proof. This is easily proved by induction so let $\mathcal{S} = \{n \in \mathbb{N} | \overline{z^n} = \bar{z}^n\}$ then we have

$1 \in \mathcal{S}$. from $\overline{z^1} = \bar{z} = \bar{z}^1$ it follows that $1 \in \mathcal{S}$

$n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$. we have

$$\begin{aligned} \overline{z^{n+1}} &= \overline{z^n \cdot z} \\ &\underset{9.5}{=} \overline{z^n} \cdot \bar{z} \\ &\underset{n \in \mathcal{S}}{=} \bar{z}^n \cdot \bar{z} \\ &= \bar{z}^n \end{aligned}$$

proving that $n + 1 \in \mathcal{S}$

□

Definition 9.10. $\mathbb{R}_{\mathbb{C}} = \{(x, y) \in \mathbb{C} | y = 0\} \subseteq \mathbb{C}$

Theorem 9.11. *If $\alpha = (\alpha', 0) \in \mathbb{R}_{\mathbb{C}}$ and $z = (x, y) \in \mathbb{C}$ then $\alpha \cdot z = (\alpha' \cdot x, \alpha' \cdot y)$*

Proof. $\alpha \cdot z = (\alpha', 0) \cdot (x, y) = (\alpha' \cdot x - 0 \cdot y, \alpha' \cdot y + 0 \cdot x) = (\alpha' \cdot x, \alpha' \cdot y)$ □

Theorem 9.12. *If $x \in \mathbb{C}$ so that $x = \bar{x}$ then $x \in \mathbb{R}_{\mathbb{C}}$*

Proof. Let $x = (a, b)$ then if $x = \bar{x}$ we have $(a, b) = (a, -b) \Rightarrow b = -b \Rightarrow 2 \cdot b = 0 \Rightarrow b = 0 \Rightarrow x = (a, 0) \in \mathbb{R}_{\mathbb{C}}$ □

Theorem 9.13. $\langle \mathbb{R}_{\mathbb{C}}, +, \cdot \rangle$ *forms a sub field (see 3.31)*

Proof. We have to prove the following (using the definition of a sub field)

$\forall x, y \in \mathbb{R}_{\mathbb{C}} \vDash x + y \in \mathbb{R}_{\mathbb{C}}$. as $x, y \in \mathbb{R}_{\mathbb{C}}$ there exists $x', y' \in \mathbb{R}$ so that $x = (x', 0) \wedge y = (y', 0)$ so that $x + y = (x', 0) + (y', 0) = (x' + y', 0 + 0) = (x + x', 0) \in \mathbb{R}$

$\forall x, y \in \mathbb{R}_{\mathbb{C}} \vDash x \cdot y \in \mathbb{R}_{\mathbb{C}}$. as $x, y \in \mathbb{R}_{\mathbb{C}}$ there exists $x', y' \in \mathbb{R}$ so that $x = (x', 0) \wedge y = (y', 0)$ hence we have $x \cdot y = (x' \cdot y' - 0 \cdot 0, x' \cdot 0 + y' \cdot 0) = (x' \cdot y', 0) \in \mathbb{R}_{\mathbb{C}}$

$\forall x \in \mathbb{R}_{\mathbb{C}} \vDash -x \in \mathbb{R}_{\mathbb{C}}$. as $x \in \mathbb{R}_{\mathbb{C}}$ there exists a $x' \in \mathbb{R}$ such that $x = (x', 0)$ hence $-x = (-x', -0) = (-x', 0) \in \mathbb{R}_{\mathbb{C}}$

$0 \in \mathbb{R}_{\mathbb{C}}$. this is trivial as $0 = (0, 0) \in \mathbb{R}_{\mathbb{C}}$ (see 9.2)

$1 \in \mathbb{R}_{\mathbb{C}}$. this is trivial as $1 = (1, 0) \in \mathbb{R}_{\mathbb{C}}$ (see 9.2) □

The following theorem combined withe the above that we can identify $\mathbb{R}$ with the sub field $\mathbb{R}_{\mathbb{C}}$ of $\mathbb{C}$

Theorem 9.14. *$i_{\mathbb{R}}: \mathbb{R} \rightarrow \mathbb{R}_{\mathbb{C}}$ defined by $i_{\mathbb{R}}(x) = (x, 0)$ is a injection with $i_{\mathbb{R}}(\mathbb{R}) = \mathbb{R}_{\mathbb{C}}$ such that $i_{\mathbb{R}}: \mathbb{R} \rightarrow \mathbb{R}_{\mathbb{C}}$ is a field isomorphism (see 3.33)*

Proof. First $i_{\mathbb{R}}(x) = i_{\mathbb{R}}(y)$ then $(x, 0) = (y, 0) \Rightarrow x = y$. Second if $x \in \mathbb{R}_{\mathbb{C}}$ then $\exists x' \in \mathbb{R}$ such that $x = (x', 0) = i_{\mathbb{R}}(x')$ proving that $i_{\mathbb{R}}: \mathbb{R} \rightarrow \mathbb{R}_{\mathbb{C}}$ is a bijection. To prove that $i_{\mathbb{R}}$ is a field isomorphism we must prove that it conserve the addition and multiplication operations.

1. If $x, y \in \mathbb{R}$ then $i_{\mathbb{R}}(x) + i_{\mathbb{R}}(y) = (x, 0) + (y, 0) = (x + y, 0 + 0) = (x + y, 0) = i_{\mathbb{R}}(x + y)$
2. If $x, y \in \mathbb{R}$ then $i_{\mathbb{R}}(x) \cdot i_{\mathbb{R}}(y) = (x, 0) \cdot (y, 0) = (x \cdot y - 0 \cdot 0, x \cdot 0 + y \cdot 0) = (x \cdot y, 0) = i_{\mathbb{R}}(x \cdot y)$
3. $i_{\mathbb{R}}(0) = (0, 0)$ the neutral element in $\mathbb{R}_{\mathbb{C}}$ and $\mathbb{C}$
4. $i_{\mathbb{R}}(1) = (1, 0)$ the neutral multiplicative element in $\mathbb{R}_{\mathbb{C}}$ and $\mathbb{C}$ □

Definition 9.15. *$i \in \mathbb{C}$ is defined by $i = (0, 1)$ and $1 \in \mathbb{R}_{\mathbb{C}}$ is defined by $1 = (1, 0)$*

Note 9.16. As $\bar{i} = \overline{(0, 1)} = (1, 0) = 1$

Definition 9.17. *If $z = (x, y) \in \mathbb{C}$ we have $\mathrm{Re}(z) = (x, 0) \in \mathbb{R}_{\mathbb{C}}$ and $\mathrm{Img}(z) = (y, 0) \in \mathbb{R}_{\mathbb{C}}$*

Lemma 9.18. *Let $z \in \mathbb{C}$ then $z = \mathrm{Re}(z) + i \cdot \mathrm{Img}(z)$ further if $z = x + i \cdot y$ where $x, y \in \mathbb{R}_{\mathbb{C}}$ then $x = \mathrm{Re}(z) \wedge y = \mathrm{Img}(z)$*

Proof. Take $z \in \mathbb{C}$ then we have $z = (z, y)$ further $i \cdot \mathrm{Img}(z) = (0, 1) \cdot (y, 0) = (0 \cdot y - 1 \cdot 0, 0 \cdot 0 + 1 \cdot y) = (0, y)$ hence $\mathrm{Re}(z) + i \cdot \mathrm{Img}(z) = (x, 0) + (0, y) = (x, y) = z$. Finally if $z = x + i \cdot y$ where $x, y \in \mathbb{R}_{\mathbb{C}}$ then $\exists x', y' \in \mathbb{R}$ such that $x = (x', 0) \wedge y = (y', 0)$ so $z = (x', 0) + i \cdot (y', 0) = (x', 0) + (0, 1) \cdot (y', 0) = (x', 0) + (0 \cdot y' - 1 \cdot 0, 0 \cdot 0 + 1 \cdot y') = (x', 0) + (0, y') = (x', y')$ hence $\mathrm{Re}(z) = (x', 0) = x$ and $\mathrm{Img}(z) = (y', 0) = y$ □

Theorem 9.19. *Let $z \in \mathbb{C}$ then $2 \cdot \mathrm{Re}(z) = z + \bar{z}$ and $2 \cdot \mathrm{Img}(z) = z - \bar{z}$*

Proof. Let $z=(x,y)$ then $z+\bar{z}=(x,y)+(x,-y)=(2\cdot x,0)=2\cdot(x,0)=2\cdot\mathrm{Re}(z)$, further $z-\bar{z}=(x,y)-(x,-y)=(0,2\cdot y)=2\cdot(0,y)=2\cdot\mathrm{Img}(z)$ □

The above motivates the following notation of complex numbers

Notation 9.20. *Let $z\in\mathbb{C}$ then we can write z in a unique way as $z=x+i\cdot y$ where $x=\mathrm{Re}(z), y=\mathrm{Img}(z)\in\mathbb{R}_{\mathbb{C}}$*

Theorem 9.21. $i^2=-1$

Proof. $i^2=(0,1)\cdot(0,1)=(0\cdot 0-1\cdot 1,0\cdot 1+1\cdot 0)=(-1,0)=-(1,0)=-1$ (the unit in $\mathbb{C}$) □

Corollary 9.22. *If $n\in\mathbb{N}_0$ then $i^{4\cdot n}=1$, $i^{4\cdot n+1}=i$, $i^{4\cdot n+2}=-1$, $i^{4\cdot n+3}=-i$*

Proof. We prove the first part by induction so let $\mathcal{S}=\{k\in\mathbb{N}_0|i^{4\cdot n}=1\}$ then we have

***n* = 0.** as $n=0$ we have $i^{4\cdot 0}=i^0=1$ and $i^{4\cdot 0+1}=i^1=i$

$n\in\mathcal{S}\Rightarrow n+1$. we have

$$\begin{aligned} i^{4\cdot(n+1)} &= i^{4\cdot n+4} \\ &= i^{4\cdot n}\cdot i\cdot i\cdot i\cdot i \\ &= i^{4\cdot n}\cdot(-1)\cdot(-1) \\ &= i^{4\cdot n} \\ &\underset{n\in\mathcal{S}}{=} 1 \end{aligned}$$

proving that $n+1$.

Mathematical induction proves then that $i^{4\cdot n}=1$. Next $i^{4\cdot n+1}=i^{4\cdot n}\cdot i=i$, $i^{4\cdot n+2}=i^{4\cdot n+1}\cdot i=i\cdot i=-1$ and finally $i^{4\cdot n+3}=i^{4\cdot n+2}\cdot i=(-1)\cdot i=-i$ □

9.2 Order relation on $\mathbb{C}$

Notation 9.23. *To avoid any confusion we use $\leqslant$ for the order relation on $\mathbb{R}$. $\leqslant_{\mathbb{C}}$ is the order relation on $\mathbb{C}$ and $\leqslant_{\mathbb{R}}$ will be the order relation on $\mathbb{R}_{\mathbb{C}}$*

We use lexical ordering (see 2.141) to define a partial order on $\mathbb{C}$ that is fully ordered.

Definition 9.24. *Define $\leqslant_{\mathbb{C}}\in\mathbb{C}\times\mathbb{C}$ by $(x,y)\leqslant_{\mathbb{C}}(x',y')$ iff either*

1. $x<x'$
2. $x=x'$ *and* $y\leqslant y'$

Using 2.141 together with 8.38 we have that $\leqslant$ is a partial order relation on $\mathbb{C}$ that is fully ordered

As $\mathbb{R}_{\mathbb{C}}\subseteq\mathbb{C}$ we can use the above and 2.139 that

Theorem 9.25. *If $x, y \in \mathbb{R}$ with $x \leqslant y \Leftrightarrow i_{\mathbb{R}}(x) \leqslant_{\mathbb{C}} i_{\mathbb{R}}(y)$ and if $x < y \Leftrightarrow i_{\mathbb{R}}(x) <_{\mathbb{C}} i_{\mathbb{R}}(y)$*

Proof. We have the following cases to consider for $x \leqslant y$

$\boldsymbol{x < y}$**.** then $i_{\mathbb{R}}(x) = (x, 0) \leqslant_{\mathbb{C}} (y, 0) = i_{\mathbb{R}}(y)$ and as $(x, 0) \neq (y, 0)$ [if $(x, 0) = (y, 0) \Rightarrow x = y$] we have $i_{\mathbb{R}}(x) < i_{\mathbb{R}}(y)$

$\boldsymbol{x = y}$**.** then as $0 \leqslant 0$ we have $(x, 0) = i_{\mathbb{R}}(x) \leqslant_{\mathbb{C}} i_{\mathbb{R}}(y) = (y, 0)$

For the opposite implication if $i_{\mathbb{R}}(x) \leqslant_{\mathbb{C}} i_{\mathbb{R}}(y)$ then we have either

$\boldsymbol{i_{\mathbb{R}}(x) < i_{\mathbb{R}}(y)}$**.** assume that $x \nless y$ then by 9.41 we have $y \leqslant x$ hence by the first part of the theorem we have $i_{\mathbb{R}}(y) \leqslant_{\mathbb{C}} i_{\mathbb{R}}(x) <_{\mathbb{C}} i_{\mathbb{R}}(y)$ giving the contradiction $i_{\mathbb{R}}(y) <_{\mathbb{C}} i_{\mathbb{R}}(y)$. Hence we must have $x < y$

$\boldsymbol{i_{\mathbb{R}}(x) = i_{\mathbb{R}}(y)}$**.** as $i_{\mathbb{R}}$ is a bijection we conclude then that $x = y$ □

Definition 9.26. $\leqslant_{\mathbb{R}} \subseteq \mathbb{R}_{\mathbb{C}} \times \mathbb{R}_{\mathbb{C}}$ *is defined by* $\{((x, 0), (y, 0)) | x, y \in \mathbb{R} \wedge x \leqslant y\}$

Theorem 9.27. $\leqslant_{\mathbb{R}} = \leqslant_{\mathbb{C}} \bigcap (\mathbb{R}_{\mathbb{C}} \times \mathbb{R}_{\mathbb{C}})$ *hence using 2.145 we have that* $\langle \mathbb{R}_{\mathbb{C}}, \leqslant_{\mathbb{R}} \rangle$ *is a fully ordered set.*

Proof. If $(z, z') \in \mathbb{R}_{\mathbb{C}}$ there exists $x, x' \in \mathbb{R}$ such that $z = (x, 0) \wedge z' = (x', 0)$, hence if $(z, z') \in \leqslant_{\mathbb{R}}$ we have $x \leqslant x'$ which proves that $(z, z') \in \leqslant_{\mathbb{C}}$, also if $(z, z') \in \leqslant_{\mathbb{C}}$ we must have $x \leqslant x'$ proving that $(z, z') \in \leqslant_{\mathbb{R}}$. So we conclude that (see 2.139)

$$\leqslant_{\mathbb{R}} = \leqslant_{\mathbb{C}} \bigcap \mathbb{R}_{\mathbb{C}} \times \mathbb{R}_{\mathbb{C}}$$

□

Corollary 9.28. $i_{\mathbb{R}}: \mathbb{R} \rightarrow \mathbb{R}_{\mathbb{C}}$ *is a order preserving field isomorphism (see isomorphism and 3.33)*

Proof. This follows from 9.14, 9.27 and 9.25 □

Theorem 9.29. *If $x, y \in \mathbb{R}_{\mathbb{C}}$ with $0 \leqslant_{\mathbb{R}} x, y$ and $x + y = 0$ then $x = y = 0$*

Proof. First there exists a $x', y' \in \mathbb{R}$ such that $x = i_{\mathbb{R}}(x') = (x', 0)$, $y = i_{\mathbb{R}}(y') = (y, 0)$ and as $0 = (0, 0) = x + y = (x', 0) + (y', 0) = (x' + y', 0)$ we have $x' + y' = 0$. Also from $0 \leqslant_{\mathbb{R}} x$ we have $0 \leqslant x'$ [if $x' < 0$ then $0 = i_{\mathbb{R}}(x') <_{\mathbb{R}} i_{\mathbb{R}}(0) = x \Rightarrow x <_{\mathbb{R}} 0 \leqslant_{\mathbb{R}} x \Rightarrow x <_{\mathbb{R}} x$ a contradiction] and similar we have $0 \leqslant y'$. So using 9.42 we have then $x' = 0 = y'$ but this means that $x = y = (0, 0) = 0$ □

Definition 9.30. $\mathbb{R}_{\mathbb{C}+} = \{\alpha \in \mathbb{R}_{\mathbb{C}} | \alpha >_{\mathbb{R}} 0\}$ *and* $\mathbb{R}_{\mathbb{C}-} = \{\alpha \in \mathbb{R}_{\mathbb{C}} | \alpha <_{\mathbb{R}} 0\}$

Based on the properties (see 9.410) of the reals and 9.13 we have the following theorem

Theorem 9.31. *We have the following for the set of reals embedded in the complex*

1. *If $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$ then we have the following exclusive possibilities*
   a. $\alpha <_{\mathbb{R}} \beta$
   b. $\beta <_{\mathbb{R}} \alpha$

c. $\alpha = \beta$

2. *If* $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$ *with* $\alpha <_{\mathbb{R}} \beta \Rightarrow -\beta <_{\mathbb{R}} -\alpha$
3. *If* $\alpha, \beta, \gamma \in \mathbb{R}_{\mathbb{C}}$ *with* $\alpha <_{\mathbb{R}} \beta \Rightarrow \alpha + \gamma <_{\mathbb{R}} \beta + \gamma$
4. *If* $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$ *and* $\alpha <_{\mathbb{R}} \beta$ *with* $\gamma \in \mathbb{R}_{\mathbb{C}+}$ *then* $\alpha \cdot \gamma <_{\mathbb{R}} \beta \cdot \gamma$
5. *If* $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$ *and* $\alpha <_{\mathbb{R}} \beta$ *with* $\gamma \in \mathbb{R}_{\mathbb{C}-}$ *then* $\beta \cdot \gamma <_{\mathbb{R}} \alpha \cdot \gamma$
6. *If* $0 <_{\mathbb{R}} \alpha \Rightarrow 0 <_{\mathbb{R}} \alpha^{-1}$
7. *If* $0 <_{\mathbb{R}} \alpha <_{\mathbb{R}} \beta \Rightarrow \beta^{-1} <_{\mathbb{R}} \alpha^{-1}$
8. *If* $\alpha \in \mathbb{R}_{\mathbb{C}} \Rightarrow 0 \leqslant_{\mathbb{R}} \alpha^2$ *and if* $\alpha \neq 0$ *then* $0 <_{\mathbb{R}} \alpha^2$
9. *If* $\alpha, \beta \in \mathbb{R}_{\mathbb{C}+} \bigcup \{0\}$ *are such that* $\alpha <_{\mathbb{R}} \beta$ *then* $\alpha^2 <_{\mathbb{R}} \beta^2$ *(so* ^{2}: $\{x \in \mathbb{R}_{\mathbb{C}} | 0 \leqslant_{\mathbb{R}} x\} \to \{x \in \mathbb{R}_{\mathbb{C}} | 0 \leqslant_{\mathbb{R}} x\}$ *is a strictly increasing function)*
10. *If* $\alpha = (\alpha', 0) \in \mathbb{R}_{\mathbb{C}+}$ *and* $n \in \mathbb{N}_0$ *then* $\alpha^n = (\alpha'^n, 0) \Rightarrow \alpha^n \in \mathbb{R}_{\mathbb{C}+}$

Proof. First note that if $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$ then there exists a α', β' such that $\alpha = (\alpha', 0)$, $\beta = (\beta', 0)$ and $\alpha \leqslant_{\mathbb{R}} \beta$ if and only if $\alpha' \leqslant \beta'$ also $\alpha = \beta$ if and only if $\alpha' = \beta'$ so that we have also $\alpha <_{\mathbb{R}} \beta$ if and only if $\alpha' < \beta'$

1. If $\alpha = (\alpha', 0), \beta = (b', 0)$ then as for α', β' we have the following possibilities

   a. $\alpha' < \beta'$

   b. $\beta' = \alpha'$

   c. $\alpha' = \beta'$

   we have the following possibilities for α, β

   a. $\alpha <_{\mathbb{R}} \beta$

   b. $\beta <_{\mathbb{R}} \alpha$

   c. $\alpha = \beta$

2. If $(\alpha', 0) = \alpha <_{\mathbb{R}} \beta = (\beta', 0) \Rightarrow \alpha' < \beta' \Rightarrow -\beta' < -\alpha' \Rightarrow -\alpha <_{\mathbb{R}} -\beta$
3. If $(\alpha', 0) = \alpha <_{\mathbb{R}} \beta = (\beta', 0) \Rightarrow \alpha' < \beta'$ and $\gamma = (\gamma', 0)$ then $\alpha' + \gamma' < \beta' + \gamma' \Rightarrow \alpha + \gamma = (\alpha' + \gamma', 0) <_{\mathbb{R}} (\beta' + \gamma', 0) = \beta + \gamma$
4. If $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$ and $\gamma \in \mathbb{R}_{\mathbb{C}+}$ then $\alpha = (\alpha', 0), \beta = (\beta', 0), \gamma = (\gamma', 0)$ where α, $\beta \in \mathbb{R}$ and $\gamma \in \mathbb{R}_+$ then from $\alpha <_{\mathbb{R}} \beta$ it follows that $\alpha' < \beta' \underset{\gamma' \in \mathbb{R}_+}{\Rightarrow} \alpha' \cdot \gamma' < \beta' \cdot \gamma' \Rightarrow \alpha \cdot \gamma = (\alpha' \cdot \gamma', 0) <_{\mathbb{R}} (\beta' \cdot \gamma', 0) = \beta \cdot \gamma$
5. If $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$ and $\gamma \in \mathbb{R}_{\mathbb{C}-}$ then $\alpha = (\alpha', 0), \beta = (\beta', 0), \gamma = (\gamma', 0)$ where α, $\beta \in \mathbb{R}$ and $\gamma \in \mathbb{R}_-$ then from $\alpha <_{\mathbb{R}} \beta$ it follows that $\alpha' < \beta' \underset{\gamma' \in \mathbb{R}_-}{\Rightarrow} \beta' \cdot \gamma' < \alpha' \cdot \gamma' \Rightarrow \beta \cdot \gamma = (\beta' \cdot \gamma', 0) <_{\mathbb{R}} (\alpha' \cdot \gamma', 0) = \alpha \cdot \gamma$
6. If $0 = (0, 0) <_{\mathbb{R}} \alpha = (\alpha', 0) \Rightarrow 0 < \alpha' \Rightarrow 0 < {\alpha'}^{-1} \Rightarrow 0 = (0, 0) <_{\mathbb{R}} ({\alpha'}^{-1}, 0) = \alpha^{-1}$
7. If $0 <_{\mathbb{R}} (\alpha', 0) = \alpha <_{\mathbb{R}} \beta = (\beta', 0) \Rightarrow 0 < \alpha' < \beta' => {\beta'}^{-1} < {\alpha'}^{-1} \Rightarrow \beta^{-1} = ({\beta'}^{-1}, 0) <_{\mathbb{R}} ({\alpha'}^{-1}, 0) < \alpha^{-1}$
8. If $\alpha = (\alpha', 0) \in \mathbb{R}_{\mathbb{C}}$ then $0 \leqslant_{\mathbb{R}} (<_{\mathbb{R}}) {\alpha'}^2$ ($<$ if $\alpha' \neq 0$) so we have $0 \leqslant_{\mathbb{R}} (<_{\mathbb{R}}) ({\alpha'}^2, 0) = \alpha^2$ ($<_{\mathbb{R}}$ if $\alpha = (\alpha', 0) \neq (0, 0) = 0$)

9. If $\alpha = (\alpha', 0), \beta = (\beta', 0) \in \mathbb{R}_{\mathbb{C}+} \bigcup \{0\}$ are such that $\alpha <_{\mathbb{R}} \beta$ then $\alpha', \beta' \in \mathbb{R}_+ \bigcup \{0\}$ so $\alpha'^2 < \beta'^2 \Rightarrow \alpha^2 = (\alpha'^2, 0) <_{\mathbb{R}} (\beta'^2, 0) = \beta^2$
10. We prove this by induction, so let $S = \{n \in \mathbb{N}_0 | \text{if} (\alpha', 0) = \alpha \in \mathbb{R}_{\mathbb{C}+} \text{then} \, \alpha^n = (\alpha'^n, 0)\}$ then we have
    a. if $n = 0$ then by definition we have $\alpha^0 = (\alpha', 0) = (\alpha'^0, 0) \Rightarrow 0 \in S$
    b. If $n \in S$ then $\alpha^{n+1} = \alpha^n \cdot a = (\alpha'^n, 0) \cdot (\alpha', 0) = (\alpha'^{n+1}, 0) \Rightarrow n + 1 \in S$ □

9.3 Norm on $\mathbb{C}$

Definition 9.32. *Given* $(x, y) \in \mathbb{C}$ *we define* $|(x, y)| = \left(\sqrt{x^2 + y^2}, 0\right) \in \mathbb{R}_{\mathbb{C}}$ *(which is defined as* $0 \leqslant x^2 + y^2 \in \mathbb{R}$*) and this is called the* ***complex norm***

Theorem 9.33. $\forall z, z' \in \mathbb{C}$ *we have*

1. $0 \leqslant_{\mathbb{R}} |z|$
2. $|z \cdot z'| = |z| \cdot |z'|$
3. $|\bar{z}| = |z|$
4. $\text{Re}(z) \leqslant_{\mathbb{R}} |z|$
5. $z \cdot \bar{z} = |z|^2$
6. *If* $x, y \in \mathbb{R}_{\mathbb{C}}$ *with* $0 \leqslant_{\mathbb{R}} x, y$ *then from* $x^2 \leqslant_{\mathbb{R}} y^2$ *we have* $x \leqslant_{\mathbb{R}} y$
7. $|z + z'| \leqslant_{\mathbb{R}} |z| + |z'|$
8. $|z| = 0 \Leftrightarrow z = 0$
9. *If* $z \in \mathbb{R}_{\mathbb{C}}$ *then* $|z| = |\text{Re}(z)|$

Proof.

1. Let $z = (x, y)$ then as $\sqrt{\ }: \{x \in \mathbb{R} | 0 \leqslant x\} \rightarrow \{x \in \mathbb{R} | 0 \leqslant x\}$ we have that $0 \leqslant_{\mathbb{R}} \sqrt{x^2 + y^2}$ given then two cases to consider:

   $\mathbf{0 < \sqrt{x^2 + y^2}}$**.** then $0 <_{\mathbb{R}} \left(\sqrt{x^2 + y^2}, 0\right) = |z|$ proving $0 \leqslant_{\mathbb{R}} |z|$

   $\mathbf{0 = \sqrt{x^2 + y^2}}$**.** then $0 = (0, 0) = \left(\sqrt{x^2 + y^2}, 0\right) = |z|$ proving $0 \leqslant_{\mathbb{R}} |z|$

2. If $z = (x, y) \in \mathbb{C}, z' = (x', y') \in \mathbb{C}$ then $z \cdot z' = (x \cdot x' - y \cdot y', x \cdot y' + y \cdot x')$ and $|z \cdot z'| = \left(\sqrt{(x \cdot x' - y \cdot y')^2 + (x \cdot y' + y \cdot x')^2}, 0\right) = \left(\sqrt{x^2 \cdot x'^2 + y^2 \cdot y'^2 - 2 \cdot x \cdot x' \cdot y \cdot y' + x^2 \cdot y'^2 + y^2 \cdot x'^2 + 2 \cdot x \cdot x' \cdot y \cdot y'}, 0\right) = \left(\sqrt{x^2 \cdot x'^2 + y^2 \cdot y'^2 + x^2 \cdot y'^2 + y^2 \cdot x'^2}, 0\right) = \left(\sqrt{x^2 \cdot (x'^2 + y'^2) + y^2 \cdot (x'^2 + y'^2)}, 0\right) = \left(\sqrt{(x^2 + y^2) \cdot (x'^2 + y'^2)}, 0\right) \underset{0 \leqslant (x^2 + y^2), (x'^2 + y'^2)}{=} \left(\sqrt{x^2 + y^2}, 0\right) \cdot \left(\sqrt{x'^2 + y'^2}, 0\right) = |z| \cdot |z'|$

3. If $z = (x, y)$ then $|\bar{z}| = |(x, -y)| = \left(\sqrt{x^2 + (-y)^2}, 0\right) = \left(\sqrt{x^2 + y^2}, 0\right) = |z|$

4. If $z=(x,y)$ then $\mathrm{Re}(z)=(x,0)$, now as $0\leqslant y^2$ we have that $x^2\leqslant x^2+y^2 \underset{\sqrt{}\,\text{is increasing (see 9.71)}}{\Rightarrow} \sqrt{|x|^2}\leqslant\sqrt{x^2+y^2}$ hence as $x\leqslant|x|=\sqrt{|x|^2}$ we have $x\leqslant\sqrt{x^2+y^2}$ so that $\mathrm{Re}(z)=(x,0)\leqslant_{\mathbb{R}}\left(\sqrt{x^2+y^2},0\right)=|z|$
5. If $z=(x,y)$ then $z\cdot\bar{z}=(x,y)\times(x,-y)=(x\cdot x+y\cdot y,-x\cdot y+y\cdot x)=(x^2+y^2,0)\underset{0\in x^2+y^2}{=}\left(\left(\sqrt{x^2+y^2}\right)^2,0\right)=\left(\sqrt{x^2+y^2},0\right)\cdot\left(\sqrt{x^2+y^2},0\right)=|z|^2$
6. Let $z_1,z\in\mathbb{R}_{\mathbb{C}}$ then there exists $x'y'\in\mathbb{R}$ such that $x=(x',0)\wedge y=(y',0)$ hence $x^2=(x'\cdot x'-0,0,x',0+0\cdot x')=(x'^2,0)$ and $y^2=(y',0)\cdot(y',0)=(y'\cdot y'-0\cdot 0,y'\cdot 0+0\cdot y')=(y'^2,0)$ now as $0\leq x,y$ we have that $0\leqslant x',y'$ and as $\sqrt{}$ os a increasing function on 9.71 we have that $x'\leqslant y$ hence $x=(x',0)=i_{\mathbb{R}}(x')\leqslant_{\mathbb{R}_{9.28}} i_{\mathbb{R}}(y')=(y',0)=y$
7. Given $x,y\in\mathbb{C}$ then we have by (5)

$$\begin{array}{rcl}
|x+y|^2 & \underset{(5)}{=} & (x+y)\cdot\overline{(x+y)}\\
& \underset{9.5}{=} & (x+y)\cdot(\bar{x}+\bar{y})\\
& = & x\cdot\bar{x}+x\cdot\bar{y}+y\cdot\bar{x}+y\cdot\bar{y}\\
& \underset{(5)}{=} & |x|^2+|y|^2+(x\cdot\bar{y}+\overline{x\cdot\bar{y}})\\
& \underset{9.19}{=} & |x|^2+|y|^2+2\cdot\mathrm{Re}(x\cdot\bar{y})\\
& \leqslant_{\mathbb{R}(4)} & |x|^2+|y|^2+2\cdot|x\cdot\bar{y}|\\
& \underset{(2)}{=} & |x|^2+|y|^2+2\cdot|x|\cdot|y|\\
& = & (|x|+|y|)^2
\end{array}$$

   hence using (6) we have $|x+y|\leqslant_{\mathbb{R}}|x|+|y|$
8. If $z=(x,y)$ then if $z=0=(0,0)$ we have $|z|=\left(\sqrt{0^2+0^2},0\right)=\left(\sqrt{(0)},0\right)=0$. On the other side if $|z|=0$ then $\left(\sqrt{x^2+y^2},0\right)=0=(0,0)\Rightarrow\left(\sqrt{x^2+y^2}\right)=0=0^2\underset{0\leqslant x^2+y^2}{\Rightarrow}x^2+y^2=0$. If now $x\neq 0\Rightarrow 0<|x|\Rightarrow 0<|x|^2=x^2\underset{0\leqslant y^2\,\text{and}\,9.41}{\Rightarrow}0<x^2+y^2$ a contradiction. So we must have that $x=0$ and thus $0=x^2+y^2=0^2+y^2=y^2$, if $y\neq 0\Rightarrow 0<|y|\Rightarrow 0<|y|^2=y^2=0$ a contradiction so we have $y=0$ and thus $z=(0,0)=0$
9. If $z\in\mathbb{R}_{\mathbb{C}}$ then $\exists z'\in\mathbb{R}$ such that $z=(x',0)$ hence $\mathrm{Re}(z)=z$ so that $|\mathrm{Re}(z)|=|z|$ □

9.4 $\mathbb{R}$ and $\mathbb{R}_{\mathbb{C}}$

9.4.1 $\mathbb{N}_{0_{\mathbb{C}}}$, $\mathbb{Z}_{\mathbb{C}}$ and $\mathbb{Q}_{\mathbb{C}}$

Definition 9.34. *Define* $\mathbb{N}_{0_{\mathbb{C}}}\subseteq\mathbb{Z}_{\mathbb{C}}\subseteq\mathbb{Q}_{\mathbb{C}}\subseteq\mathbb{R}_{\mathbb{C}}\subseteq\mathbb{C}$ *by* $\mathbb{N}_{0_{\mathbb{C}}}=\{(x,0)\in\mathbb{R}_{\mathbb{C}}|x\in\mathbb{N}_{0_{\mathbb{R}}}\}$, $\mathbb{Z}_{\mathbb{C}}=\{(x,0)|x\in\mathbb{Z}_{\mathbb{R}}\}$, $\mathbb{Q}_{\mathbb{C}}=\{(x,0)\in\mathbb{R}_{\mathbb{C}}|x\in\mathbb{Q}_{\mathbb{R}}\}$

Theorem 9.35. *$\langle \mathbb{N}_{0\mathbb{C}}, + \rangle, \langle \mathbb{N}_{0\mathbb{C}}, \cdot \rangle$ forms a semi-groups and if we define $i_{\mathbb{R}_{\mathbb{N}_0}}$: $\mathbb{N}_0 \rightarrow \mathbb{N}_{0\mathbb{C}}$ by $i_{\mathbb{R}_{\mathbb{N}_0}} = i_{\mathbb{R}_{|\mathbb{N}_{0\mathbb{R}}}} \circ i_{\mathbb{Q}_{\mathbb{N}_0}}$ (see 8.56) then $i_{\mathbb{R}_{\mathbb{N}_0}}$ is a a order preserving semi-group isomorphism (for $+$ and $\cdot$)*

Proof. First we prove that $\langle \mathbb{N}_{0\mathbb{C}}, + \rangle$ and $\langle \mathbb{N}_{0\mathbb{C}}, \cdot \rangle$ are semi-groups. If $x, y \in \mathbb{N}_{0\mathbb{C}}$ then we have $x = (x', 0), y = (y', 0)$ and $x', y' \in \mathbb{N}_{0\mathbb{R}}$ so that

1. $x + y = (x', 0) + (y', 0) = (x' + y', 0) \in \mathbb{N}_{0\mathbb{C}}$ as $x' + y' \in \mathbb{N}_{0\mathbb{R}}$ (for $\mathbb{N}_{0\mathbb{R}}$ is a semi-group)
2. $0 = (0, 0) \in \mathbb{N}_{0\mathbb{C}}$
3. $x \cdot y = (x' \cdot y' - 0 \cdot 0, x' \cdot 0 + 0 \cdot y') = (x' \cdot y', 0) \in \mathbb{N}_{0\mathbb{C}}$ (as $x' \cdot y' \in \mathbb{N}_{0\mathbb{C}}$ which is a semi-group)
4. $1 = (1, 0) \in \mathbb{N}_{0\mathbb{C}}$

To prove that $i_{\mathbb{R}_{\mathbb{N}_0}}$ is a order preserving semi-group isomorphism, note that by 8.56 $i_{\mathbb{Q}_{\mathbb{N}_0}}$ is a order preserving semi-group isomorphism and that by 9.14,9.25 $i_{\mathbb{R}}$ is a order preserving isomorphism and thus $i_{\mathbb{R}_{|\mathbb{N}_0}}$ is a order preserving semi-group isomorphism. □

Theorem 9.36. *$\langle \mathbb{Z}_{\mathbb{C}}, +, \cdot \rangle$ forms a sub ring of $\langle \mathbb{C}, +, \cdot \rangle$ [as a field $\langle \mathbb{C}, +, \cdot \rangle$ is also a ring] with additive neutral element $0 = (0, 0)$ and multiplicative unit $(1, 0)$. Further there exists a order preserving ring isomorphism between $\mathbb{Z}$ and $\mathbb{Z}_{\mathbb{C}}$. So we can consider $\mathbb{Z}_{\mathbb{C}}$ as a embedding of $\mathbb{Z}$ in $\mathbb{C}$.*

Proof. If $x, y \in \mathbb{Z}_{\mathbb{C}}$ then there exists a $x', y' \in \mathbb{Z}_{\mathbb{R}}$ such that $x = (x', 0) \wedge y = (y', 0)$. We have then that

1. $x + y = (x', 0) + (y', 0) = (x' + y', 0) \in \mathbb{Z}_{\mathbb{C}}$ as $x' + y' \in \mathbb{Z}_{\mathbb{R}}$ (see 8.53)
2. $x \cdot y = (x', 0) \cdot (y', 0) = (x' \cdot y' - 0 \cdot 0, x' \cdot 0 + y' \cdot 0) = (x' \cdot y', 0) \in \mathbb{Z}_{\mathbb{C}}$ as $x' \cdot y' \in \mathbb{Z}_{\mathbb{R}}$ (see 8.53)
3. $-x = -(x', 0) \underset{9.2}{=} (-x', -0) = (-x', 0) \in \mathbb{Z}_{\mathbb{C}}$ as $-x' \in \mathbb{Z}_{\mathbb{R}}$ (see 8.53)
4. $1 = (1, 0) \in \mathbb{Z}_{\mathbb{C}}$
5. $0 = (0, 0) \in \mathbb{Z}_{\mathbb{C}}$

Next $i_{\mathbb{R}_{|\mathbb{Z}_{\mathbb{C}}}}: Z_{\mathbb{R}} \rightarrow \mathbb{Z}_{\mathbb{C}}$ is a order preserving field (and thus ring) isomorphism (see 9.14,9.25) hence as by 8.53 $i_{\mathbb{Q}_{\mathbb{Z}}}: \mathbb{Z} \rightarrow \mathbb{Z}_{\mathbb{R}}$ is a order preserving ring isomorphism we have that $i_{\mathbb{R}_{|\mathbb{Z}_{\mathbb{C}}}} \circ i_{\mathbb{Q}_{\mathbb{Z}}}: \mathbb{Z} \rightarrow \mathbb{Z}_{\mathbb{R}}$ is a order preserving ring isomorphism. □

Theorem 9.37. *$\langle \mathbb{Q}_{\mathbb{C}}, +, \cdot \rangle$ forms a sub field of $\langle \mathbb{C}, +, \cdot \rangle$ with additive neutral element $0 = (0, 0)$ and multiplicative unit $(1, 0)$. Further there exists a order preserving field isomorphism between $\mathbb{Q}$ and $\mathbb{Q}_{\mathbb{C}}$. So we can consider $\mathbb{Q}_{\mathbb{C}}$ as a embedding of $\mathbb{Q}$ in $\mathbb{C}$.*

Proof. If $x, y \in \mathbb{Q}_{\mathbb{C}}$ then there exists a $x', y' \in \mathbb{Q}_{\mathbb{R}}$ such that $x = (x', 0) \wedge y = (y', 0)$. We have then that

1. $x + y = (x', 0) + (y', 0) = (x' + y', 0) \in \mathbb{Q}_{\mathbb{C}}$ as $x' + y' \in \mathbb{Q}_{\mathbb{R}}$ (see 8.28)

2. $x \cdot y = (x', 0) \cdot (y', 0) = (x' \cdot y' - 0 \cdot 0, x' \cdot 0 + y' \cdot 0) = (x' \cdot y', 0) \in \mathbb{Z}_{\mathbb{C}}$ as $x' \cdot y' \in \mathbb{Z}_{\mathbb{R}}$ (see 8.28)
3. $-x = -(x', 0) \underset{9.2}{=} (-x', -0) = (-x', 0) \in \mathbb{Q}_{\mathbb{C}}$ as $-x' \in \mathbb{Q}_{\mathbb{R}}$ (see 8.28)
4. $1 = (1, 0) \in \mathbb{Q}_{\mathbb{C}}$
5. $0 = (0, 0) \in \mathbb{Q}_{\mathbb{C}}$

Next $i_{\mathbb{R}_{|\mathbb{Q}_{\mathbb{C}}}}: \mathbb{Q}_{\mathbb{R}} \to \mathbb{Q}_{\mathbb{C}}$ is a order preserving field isomorphism (see 9.14,9.25) hence as by 8.28 $i_{\mathbb{Q}}: \mathbb{Q} \to \mathbb{Q}_{\mathbb{R}}$ is a order preserving ring isomorphism we have that $i_{\mathbb{R}_{|\mathbb{Q}_{\mathbb{C}}}} \circ i_{\mathbb{Q}}: \mathbb{Q} \to \mathbb{Q}_{\mathbb{R}}$ is a order preserving ring isomorphism. □

Theorem 9.38. *$\mathbb{N}_{0_{\mathbb{C}}}, \mathbb{Z}_{\mathbb{C}}$ and $\mathbb{Q}_{\mathbb{C}}$ are denumerable.*

Proof. This follows from the fact that $\mathbb{N}_0$, $\mathbb{Z}, \mathbb{Q}$ are denumerable and isomorphic (and thus bijective) with $\mathbb{N}_{0_{\mathbb{C}}}$, $Z_{\mathbb{C}}$ and $\mathbb{Q}_{\mathbb{C}}$. □

Next we prove that $\mathbb{C}$ is isomorphic to $\mathbb{R}_{\mathbb{C}}^2$

Theorem 9.39. *The map $\mathcal{C}: \mathbb{R}_{\mathbb{C}} \times \mathbb{R}_{\mathbb{C}} \to \mathbb{C}$ defined by $\mathcal{C}((x, y)) = x + i \cdot y$ is a bijection.*

Proof. First we prove that C is a bijection

injectivity. Assume that $\mathcal{C}((x_1, y_1)) = \mathcal{C}((x_2, y_2))$ then there exists $x_1', y_1', x_2', y_2' \in \mathbb{R}$ such that $x_1 = (x_1', 0)$, $x_2 = (x_2', 0)$, $y_1 = (y_1', 0)$ and $y_2 = (y_2', 0)$ hence $(x_1', y_1') = (x', 0) + (0, y_1') = (x', 0) + (0, 1) \cdot (y_1', 0) = x_1 + i \cdot y_1 = \mathcal{C}((x_1, y_1)) = \mathcal{C}((x_2, y_2)) = x_2 + i \cdot y_2 = (x_2', 0) + (0, 1) \cdot (y_2', 0) = (x_2', 0) + (0, y_2') = (x_2', y_2')$ proving that $x_1' = x_2' \wedge y_1' = y_2' \Rightarrow x_1 = x_2 \wedge y_1 = y_2 \Rightarrow (x_1, y_1) = (x_2, y_2)$ proving injectivity.

surjectivity. Take $z \in \mathbb{C}$ then there exists $x', y' \in \mathbb{R}$ such that $z = (x', y') = (x', 0) + (0, y') = (x', 0) + (0, 1) \cdot (y', 0) = x + i \cdot y$ where $x = (x', 0), y = (y', 0) \in \mathbb{R}_{\mathbb{C}}$ hence $z = \mathcal{C}((x, y))$ proving that $\mathcal{C}$ is bijective. □

9.4.2 Properties of $\mathbb{R}_{\mathbb{C}}$

We show now that $\mathbb{R}$ and $\mathbb{R}_{\mathbb{C}}$ has the same properties

Definition 9.40. $\mathbb{R}_{\mathbb{C}+} = \{z \in \mathbb{R}_{\mathbb{C}} | 0 <_{\mathbb{R}} z\}$ *and* $\mathbb{R}_{\mathbb{C}-} = \{z \in \mathbb{R}_{\mathbb{C}} | z <_{\mathbb{R}} 0\}$

Theorem 9.41. *We have the following for the set of the reals*

1. *$\alpha \in \mathbb{R}_{\mathbb{C}_+}$ iff $0 <_{\mathbb{R}} \alpha$*
2. *$\alpha \in \mathbb{R}_{\mathbb{C}_-}$ iff $\alpha <_{\mathbb{R}} 0$*
3. *If $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$ then we have the following exclusive possibilities*
    a. *$\alpha <_{\mathbb{R}} \beta$*
    b. *$\beta <_{\mathbb{R}} \alpha$*

c. $\alpha = \beta$

4. *If* $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$ *with* $\alpha <_{\mathbb{R}} \beta \Rightarrow -\beta <_{\mathbb{R}} -\alpha$
5. *If* $\alpha, \beta, \gamma \in \mathbb{R}_{\mathbb{C}}$ *with* $\alpha <_{\mathbb{R}} \beta \Rightarrow \alpha + \gamma <_{\mathbb{R}} \beta + \gamma$
6. *If* $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$ *and* $\alpha <_{\mathbb{R}} \beta$ *with* $\gamma \in \mathbb{R}_{\mathbb{C}_+}$ *then* $\alpha \cdot \gamma <_{\mathbb{R}} \beta \cdot \gamma$ *[as we have* $\alpha = \beta \Rightarrow \alpha \cdot \lambda = \beta \cdot \lambda$ *we have also that* $\alpha \leqslant \beta$ *implies* $\alpha \cdot \lambda \leqslant_{\mathbb{R}} \beta \cdot \lambda$*]*
7. *If* $\alpha, \beta, \gamma, \delta \in \mathbb{R}_{\mathbb{C}_+}$ *with* $\alpha < \beta \wedge \gamma < \delta$ *then* $\alpha \cdot \gamma < \beta \cdot \delta$
8. *If* $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$ *and* $\alpha <_{\mathbb{R}} \beta$ *with* $\gamma \in \mathbb{R}_{\mathbb{C}_-}$ *then* $\beta \cdot \gamma <_{\mathbb{R}} \alpha \cdot \gamma$ *[as we have* $\alpha = \beta \Rightarrow \alpha \cdot \lambda = \beta \cdot \lambda$ *we have also that* $\alpha \leqslant_{\mathbb{R}} \beta$ *implies* $\beta \cdot \lambda \leqslant_{\mathbb{R}} \alpha \cdot \lambda$*]*
9. *If* $0 <_{\mathbb{R}} x \Rightarrow 0 <_{\mathbb{R}} x^{-1}$
10. *If* $0 <_{\mathbb{R}} x <_{\mathbb{R}} y \Rightarrow y^{-1} <_{\mathbb{R}} x^{-1}$
11. *If* $x \in \mathbb{R}_{\mathbb{C}} \Rightarrow 0 \leqslant_{\mathbb{R}} x^2$ *and if* $x \neq 0$ *then* $0 <_{\mathbb{R}} x^2$
12. *If* $x, y \in \mathbb{R}_{\mathbb{C}_+} \bigcup \{0\}$ *are such that* $x <_{\mathbb{R}} y$ *then* $x^2 <_{\mathbb{R}} y^2$ *(so* $^2: \{x \in \mathbb{R}_{\mathbb{C}} | 0 \leqslant_{\mathbb{R}} x\} \rightarrow \{x \in \mathbb{R}_{\mathbb{R}} | 0 \leqslant x\}$ *is a strictly increasing function)*
13. *If* $\alpha \in \mathbb{R}_{\mathbb{C}+}$ *and* $n \in \mathbb{N}_0$ *then* $\alpha^n \in \mathbb{R}_{\mathbb{C}+}$
14. *If* $\alpha \in \mathbb{R}_{\mathbb{C}+}$ *so that* $0 <_{\mathbb{R}} \alpha <_{\mathbb{R}} 1$ *then if* $n \in \mathbb{N}$ *we have* $0 <_{\mathbb{R}} \alpha^n <_{\mathbb{R}} \alpha$
15. *If* $\alpha \in \mathbb{R}_{\mathbb{C}}$ *such that* $1 \leqslant_{\mathbb{R}} \alpha$

    a. if $\alpha <_{\mathbb{R}} \beta$ *then* $\alpha <_{\mathbb{R}} \beta^n$

    b. if $\alpha \leqslant_{\mathbb{R}} \beta$ *then* $\alpha \leqslant_{\mathbb{R}} \beta^n$

    c. if $n, m \in \mathbb{N}_0$ *with* $n \leqslant m$ *we have* $\alpha^n \leqslant \alpha^m$

Proof.

1. This follows trivially from the definition of $\mathbb{R}_{\mathbb{C}+}$
2. This follows trivially from the definition of $\mathbb{R}_{\mathbb{C}+}$
3. If $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$ then by 9.14 there exists $\alpha', \beta' \in \mathbb{R}$ so that $\alpha = i_{\mathbb{R}}(\alpha')$ and $\beta = i_{\mathbb{R}}(\beta')$ using 8.39 (3) we have then for $\alpha', \beta' \in \mathbb{R}$ the following excluding cases

    $\boldsymbol{\alpha' < \beta'}$**.** then using 9.25 we have $\alpha = i_{\mathbb{R}}(\alpha') <_{\mathbb{R}} i_{\mathbb{R}}(\beta') = \beta$

    $\boldsymbol{\beta' < \alpha'}$**.** then using 9.25 we have $\alpha = i_{\mathbb{R}}(\beta') <_{\mathbb{R}} i_{\mathbb{R}}(\alpha') = \beta$

    $\boldsymbol{\alpha' = \beta}$**'.** then using 9.14 we have $\alpha = i_{\mathbb{R}}(\alpha') = i_{\mathbb{R}}(\beta') = \beta$

4. If $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$ with $\alpha <_{\mathbb{R}} \beta$ then by 9.14 there exists $\alpha', \beta' \in \mathbb{R}$ so that $\alpha = i_{\mathbb{R}}(\alpha')$, $\beta = i_{\mathbb{R}}(\beta')$, and using 9.25 we have $\alpha' < \beta'$. Using 8.39 we have $-\beta' < -\alpha' \underset{9.25}{\Rightarrow} i_{\mathbb{R}}(-\beta') <_{\mathbb{R}} i_{\mathbb{R}}(-\alpha') \underset{3.35}{\Rightarrow} -i_{\mathbb{R}}(\beta') <_{\mathbb{R}} -i_{\mathbb{R}}(\alpha') \Rightarrow -\beta <_{\mathbb{R}} = \alpha$
5. If $\alpha, \beta, \gamma \in \mathbb{R}_{\mathbb{C}}$ then by 9.14 there exists $\alpha', \beta', \gamma' \in \mathbb{R}$ so that $\alpha = i_{\mathbb{R}}(\alpha')$, $\beta = i_{\mathbb{R}}(\beta')$, $\gamma = i_{\mathbb{R}}(\gamma')$. As $\alpha <_{\mathbb{R}} \beta$ we have $i_{\mathbb{R}}(\alpha') <_{\mathbb{R}} i_{\mathbb{R}}(\beta')$ and thus by 9.25 $\alpha' < \beta'$. Using then 8.39 (5) we have that $\alpha' + \gamma' < \beta' + \gamma'$ so that using the fact that $i_{\mathbb{R}}$ is a field isomorphism we have $\alpha + \gamma = i_{\mathbb{R}}(\alpha') + i_{\mathbb{R}}(\beta') = i_{\mathbb{R}}(\alpha' + \gamma') <_{\mathbb{R}_{9.25}} i_{\mathbb{R}}(\beta' + \gamma') = i_{\mathbb{R}}(\beta') + i_{\mathbb{R}}(\gamma') = \beta + \gamma$

6. If $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$ with $\alpha <_{\mathbb{R}} \beta$, $\lambda \in \mathbb{R}_{\mathbb{C}+}$ then by 9.14 there exists $\alpha', \beta', \lambda' \in \mathbb{R}$ so that $\alpha = i_{\mathbb{R}}(\alpha')$, $\beta = i_{\mathbb{R}}(\beta')$, $\lambda = i_{\mathbb{R}}(\lambda')$. As $\alpha <_{\mathbb{R}} \beta$ we have $i_{\mathbb{R}}(\alpha') <_{\mathbb{R}} i_{\mathbb{R}}(\beta')$ and thus by 9.25 $\alpha' < \beta'$, further as $0 = i_{\mathbb{R}}(0)$ we have also $0 < \lambda'$. Using 8.39 (6) we have then $\alpha' \cdot \lambda' < \beta' \cdot \lambda'$ and hence $\alpha \cdot \lambda = i_{\mathbb{R}}(\alpha') \cdot i_{\mathbb{R}}(\lambda') = i_{\mathbb{R}}(\alpha' \cdot \lambda') <_{\mathbb{R}} i_{\mathbb{R}}(\beta' \cdot \lambda') = i_{\mathbb{R}}(\beta') \cdot i_{\mathbb{R}}(\lambda') = \beta \cdot \lambda$.

7. If $\alpha, \beta, \gamma, \delta \in \mathbb{R}_{\mathbb{C}_+}$ with $\alpha < \beta \wedge \gamma < \delta$ then by (6) we have $\alpha \cdot \gamma < \beta \cdot \gamma$ and $\beta \cdot \gamma < \beta \cdot \delta$ so that $\alpha \cdot \gamma < \beta \cdot \delta$

8. If $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$ with $\alpha <_{\mathbb{R}} \beta$, $\lambda \in \mathbb{R}_{\mathbb{C}-}$ then by 9.14 there exists $\alpha', \beta', \lambda' \in \mathbb{R}$ so that $\alpha = i_{\mathbb{R}}(\alpha')$, $\beta = i_{\mathbb{R}}(\beta')$, $\lambda = i_{\mathbb{R}}(\lambda')$. As $\alpha <_{\mathbb{R}} \beta$ we have $i_{\mathbb{R}}(\alpha') <_{\mathbb{R}} i_{\mathbb{R}}(\beta')$ and thus by 9.25 $\alpha' < \beta'$, further as $0 = i_{\mathbb{R}}(0)$ we have also $\lambda' < 0$. Using 8.39 (7) we have then $\beta' \cdot \lambda' < \alpha' \cdot \lambda'$ and hence $\beta \cdot \lambda = i_{\mathbb{R}}(\beta') \cdot i_{\mathbb{R}}(\lambda') = i_{\mathbb{R}}(\beta' \cdot \lambda') <_{\mathbb{R}} i_{\mathbb{R}}(\alpha' \cdot \lambda') = i_{\mathbb{R}}(\alpha') \cdot i_{\mathbb{R}}(\lambda') = \alpha \cdot \lambda$.

9. If $x \in \mathbb{R}_{\mathbb{C}}$ with $0 <_{\mathbb{R}} x$ then there exists a $x' \in \mathbb{R}$ such that $i_{\mathbb{R}}(x') = x$, using 9.25 and $0 = i_{\mathbb{R}}(0)$ we have $0 < x'$. Using 8.39 (8) we have then that $0 < x'^{-1} \Rightarrow 0 = i_{\mathbb{R}}(0) <_{\mathbb{R}} i_{\mathbb{R}}(x'^{-1}) \underset{3.35}{=} (i_{\mathbb{R}}(x'))^{-1} = x^{-1}$

10. If $x, y \in \mathbb{R}_{\mathbb{C}}$ with $0 <_{\mathbb{R}} x <_{\mathbb{R}} y$ then using 9.14 and 9.25 there exists $x', y' \in \mathbb{R}$ such that $0 < x' < y'$ hence using 8.39 we have that $y'^{-1} < x'^{-1}$ and thus $y^{-1} = (i_{\mathbb{R}}(y'))^{-1} \underset{3.35}{=} i_{\mathbb{R}}(y'^{-1}) <_{\mathbb{R}} i_{\mathbb{R}}(x'^{-1}) \underset{3.35}{=} (i_{\mathbb{R}}(x'))^{-1} = x^{-1}$

11. If $x \in \mathbb{R}$ then then using 9.14 there exists a $x' \in \mathbb{R}$ such that $x = i_{\mathbb{R}}(x')$. Using 8.39 (10) we have $0 \leqslant x'^2$ and thus $0 = i_{\mathbb{R}}(0) \leqslant_{\mathbb{R}} i_{\mathbb{R}}(x'^2) = i_{\mathbb{R}}(x') \cdot i_{\mathbb{R}}(x') = x \cdot x = x^2$. Further if $x \neq 0$ then as $i_{\mathbb{R}}$ is a bijection we have that $x' \neq 0$ and thus by 8.39 (10) we have $0 < x'^2$ so that $0 = i_{\mathbb{R}}(0) <_{\mathbb{R}} i_{\mathbb{R}}(x'^2) = i_{\mathbb{R}}(x') \cdot i_{\mathbb{R}}(x') = x \cdot x = x^2$

12. Let $x, y \in \mathbb{R}_{\mathbb{C}+} \bigcup \{0\}$ are such that $x <_{\mathbb{R}} y$ we have using 9.14 and 9.25 that $\exists x', y' \in \mathbb{R}$ such that $x = i_{\mathbb{R}}(x')$, $y = i_{\mathbb{R}}(y')$ and $x' < y'$. Using 8.39 we have then $x'^2 < y'^2$ so that $x^2 = i_{\mathbb{R}}(x') \cdot i_{\mathbb{R}}(x') = i_{\mathbb{R}}(x' \cdot x') = i_{\mathbb{R}}(x'^2) <_{\mathbb{R}} i_{\mathbb{R}}(y'^2) = i_{\mathbb{R}}(y' \cdot y') = i_{\mathbb{R}}(y') \cdot i_{\mathbb{R}}(y') = y^2$

13. If $\alpha \in \mathbb{R}_{\mathbb{C}+}$ then $0 <_{\mathbb{R}} \alpha$ we proceed now by induction on n to prove that $\forall n \in \mathbb{N}_0$ we have $0 < \alpha^n$. So let $X = \{n \in \mathbb{N}_0 | \text{if}\, 0 <_{\mathbb{R}} \alpha \,\text{then}\, 0 <_{\mathbb{R}} \alpha^n\}$ then we have:

    a. $\alpha^0 = 1 >_{\mathbb{R}} 0 \Rightarrow 0 \in X$

    b. If $n \in X$ then if $0 <_{\mathbb{R}} \alpha$ we have $0 <_{\mathbb{R}} \alpha^n$. Now $\alpha^{n+1} = \alpha^n \cdot \alpha \underset{0 < \alpha^n, \alpha \in \mathbb{R}_+ \text{and}\,(6)}{\Rightarrow} 0 <_{\mathbb{R}} \alpha^n \cdot \alpha = \alpha^{n+1} \Rightarrow n+1 \in X$

    Using mathematical induction (see 4.10) we have $X = \mathbb{N}_0$ proving our assertion.

14. Let $\alpha \in \mathbb{R}_{\mathbb{C}+}$ with $0 <_{\mathbb{R}} \alpha <_{\mathbb{R}} 1$ then if $n \in \mathbb{N}$ we have by (12) already $0 <_{\mathbb{R}} \alpha^n$. Now to prove $\alpha^n <_{\mathbb{R}} \alpha$ take $S = \{n < \mathbb{N} | \text{if}\, \alpha <_{\mathbb{R}} 1 \,\text{then}\, \alpha^n <_{\mathbb{R}} \alpha\}$ then we have

    a. if $n = 1$ then $\alpha^1 = \alpha <_{\mathbb{R}} 1$ so that $1 \in S$

    b. if $n \in S$ then $\alpha^{n+1} = \alpha^n \cdot \alpha <_{\mathbb{R}} \alpha$ [as $n \in S \Rightarrow \alpha^n <_{\mathbb{R}} 1 \underset{0 < \alpha \,\text{and}\,(6)}{\Rightarrow} \alpha^n \cdot \alpha <_{\mathbb{R}} 1 \cdot \alpha = \alpha$]

15.

a. We prove this by induction so let $\mathcal{S}=\{n\in\mathbb{N}|\alpha<_{\mathbb{R}}\beta^n\}$ then we have

$1\in\mathcal{S}$. this follows from $\alpha<_{\mathbb{R}}\beta=\beta^1$

$n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$. as $n\in\mathcal{S}$ we have $\alpha<_{\mathbb{R}}\beta^n$ we have by (6) that $\alpha\cdot\beta<_{\mathbb{R}}\beta\cdot\beta^n=\beta^{n+1}$ and $1\leq_{\mathbb{R}}\alpha\Rightarrow\beta\leqslant_{\mathbb{R}}\alpha\cdot\beta\underset{\alpha<\beta}{\Rightarrow}\alpha<_{\mathbb{R}}\alpha\cdot\beta<_{\mathbb{R}}\beta^{n+1}$ proving that $n+1\in\mathcal{S}$

b. We prove this by induction so let $\mathcal{S}=\{n\in\mathbb{N}|\alpha\leqslant_{\mathbb{R}}\beta^n\}$ then we have

$1\in\mathcal{S}$. this follows from $\alpha\leqslant_{\mathbb{R}}\beta=\beta^1$

$n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$. as $n\in\mathcal{S}$ we have $\alpha\leqslant\beta^n$ we have by (6) that $\alpha\cdot\beta\leqslant_{\mathbb{R}}\beta\cdot\beta^n=\beta^{n+1}$ and $1\leq_{\mathbb{R}}\alpha\Rightarrow\beta\leqslant_{\mathbb{R}}\alpha\cdot\beta\underset{\alpha\leqslant\beta}{\Rightarrow}\alpha\leqslant_{\mathbb{R}}\alpha\cdot\beta\leqslant_{\mathbb{R}}\beta^{n+1}$ proving that $n+1\in\mathcal{S}$

c. We prove this by induction so let $n\in\mathbb{N}_0$ and $\mathcal{S}_n=\{k\in\mathbb{N}_0|\alpha^n\leqslant\alpha^{n+k}\}$ then we have

$k=0$. then $\alpha^n=\alpha^{n+0}$ proving that $0\in\mathcal{S}_n$

$k\in\mathcal{S}_n\Rightarrow k+1\in\mathcal{S}_n$. then $\alpha^n\leqslant a^{n+k}\Rightarrow a^{n+1}\leqslant a^{n+k+1}$, as $1\leqslant a$ we have $a^n\leqslant a^{n+1}$ we have $a^n\leqslant a^{n+k+1}$ proving that $k+1\in\mathcal{S}_n$

Using mathematical induction we have then $\mathcal{S}_n=\mathbb{N}_0$, so if $n,m\in\mathbb{N}_0$ with $n\leqslant m$ we have that $m-n\in\mathbb{N}_0$ and thus $\alpha^n\leqslant a^{n+(m-n)=a^m}$ □

Theorem 9.42. *If $\alpha,\beta\in\mathbb{R}_{\mathbb{C}}$ with $0\leqslant\alpha,\beta$ then from $\alpha+\beta=0$ we have $\alpha=\beta=0$*

Proof. We have to consider the following cases

$0<\alpha$. then $\beta=0+\beta<\alpha+\beta=0$ giving the contradiction $\beta<0$

$0<\beta$. then $\alpha=0+\alpha<\beta+\alpha=0$ giving the contradiction $\alpha<0$

$\alpha=0=\beta$. this is what we have to prove □

Theorem 9.43. ($\mathbb{R}_{\mathbb{C}}$ is conditional complete) $\langle\mathbb{R}_{\mathbb{C}},<+_{\mathbb{R}}\rangle$ *is conditional complete (see 2.175) Using the definition of conditional completeness this means that $\forall S\subseteq\mathbb{R}\vdash S\neq\emptyset$ for which there exists a $b\in\mathbb{R}$ such that $\forall s\in S$ we have $s\leqslant b$ (existence of a upper bound) we have the existence of $\sup(S)$ (a lowest upper bound). In other words: any nonempty set in $\mathbb{R}$ with a upper bound has a lowest upper bound. Using 2.176 we see that $\forall S\subseteq\mathbb{R}\vdash S\neq\emptyset$ for which there exists a $b\in\mathbb{R}$ such that $\forall s\in S$ we have $b\leqslant s$ (existence of a lower bound) there exists a $\inf(S)$ (in other words every non empty set in $\mathbb{R}$ with a lower bound has greatest lower bound).*

Proof. As by 9.28 $i_{\mathbb{R}}:\mathbb{R}\rightarrow\mathbb{R}_{\mathbb{C}}$ is a order preserving isomorphism, $\langle R,\leqslant\rangle$ is conditionally complete we have by 2.178 that $\langle\mathbb{R}_{\mathbb{C}},\leqslant_{\mathbb{R}}\rangle$ is conditionally complete. □

Theorem 9.44. *Let $S\subseteq\mathbb{R}_{\mathbb{C}}$ then we have for $-S=\{-s,s\in S\}$ that*

1. *If $\sup(S)$ exists then $\inf(-S)$ exists and $\inf(-S)=-\sup(S)$*

2. *If* $\inf(S)$ *exists then* $\sup(-S)$ *exists and* $\sup(-S) = -\inf(S)$

Proof.

1. As $\sup(S)$ exists we have $\forall s \in -S$ that $-s \in S$ so that $-s \leqslant \sup(S)$ or $-\sup(S) \leqslant s$ so $-S$ is bounded below by $-\sup(S)$ proving by the conditional completeness of $\mathbb{R}$ [see 8.43] that

$$\inf(-S) \text{ exists and } -\sup(S) \leqslant \inf(-S). \tag{9.1}$$

Assume now that $-\sup(S) < \inf(-S)$ then $-\inf(-S) < \sup(S)$ so there exist a $s \in S$ such that $-\inf(-S) < s \leqslant \sup(S)$ so that $-s < \inf(-S)$ which as $-s \in -S$ gives $-s < \inf(-S) \leqslant -s$ a contradicton. So we must have that $\inf(-S) \leqslant -\sup(S)$ which together with (9.1) proves that

$$\inf(-S) = -\sup(S)$$

2. As $\inf(S)$ exists we have $\forall s \in -S$ that $-s \in S$ so that $\inf(S) \leqslant -s$ or $s \leqslant -\inf(S)$ so $-S$ is bounded above by $-\inf(S)$ proving by the conditional completeness of $\mathbb{R}$ [see 8.43] that

$$\sup(-S) \text{ exists and } \sup(-S) \leqslant -\inf(S) \tag{9.2}$$

Assume now that $\sup(-S) < -\inf(S)$ then $\inf(S) < -\sup(-S)$ so there exist a $s \in S$ such that $\inf(S) \leqslant s < -\sup(-S)$ so that $\sup(-S) < -s$ which as $-s \in -S$ gives $-s \leqslant \sup(-S) < -s$ a contradicton. So we must have that $-\inf(S) \leqslant \sup(-S)$ which together with (9.2) proves that

$$\sup(-S) = -\inf(S)$$ □

Theorem 9.45. *If* $S \subseteq \mathbb{R}_{\mathbb{C}}$ *has a supremum and* $\alpha \in \mathbb{R}_{\mathbb{C}}$ *with* $\alpha \geqslant 0$ *then* $\alpha \cdot S = \{\alpha \cdot s | s \in S\}$ *has a supremum equal to* $\alpha \cdot \sup(S)$

Proof. If $\alpha = 0$ then $\alpha \cdot S = \{0\}$ with $\sup(\alpha \cdot S) = \sup(\{0\}) = 0 = 0 \cdot \sup(S)$ so we only have to prove the case that $\alpha > 0$. By the definition of $\sup(S)$ we have that $\forall s \in S$ that $s \leqslant_{\mathbb{R}} \sup(S)$ so if $y \in \alpha \cdot S$ then $y = \alpha \cdot s \leqslant_{\mathbb{R}} \alpha \cdot \sup(S)$ [as $\alpha >_{\mathbb{R}} 0$] proving that $\alpha \cdot \sup(S)$ is a upper bound for $\alpha \cdot S$. By the above theorem $\alpha \cdot S$ has a $\sup(\alpha \cdot S)$ with by definition $\sup(\alpha \cdot S) \leqslant_{\mathbb{R}} \alpha \cdot \sup(S)$. If now $\sup(\alpha \cdot S) <_{\mathbb{R}} \alpha \cdot \sup(S)$ then $\frac{1}{\alpha} \cdot \sup(\alpha \cdot S) <_{\mathbb{R}} \sup(S)$ and there exists a $s \in S$ with $\frac{1}{\alpha} \cdot \sup(\alpha \cdot S) <_{\mathbb{R}} s \leqslant_{\mathbb{R}} \sup(S)$ giving $\sup(\alpha \cdot S) <_{\mathbb{R}} \alpha \cdot s \leqslant_{\mathbb{R}} \alpha \cdot \sup(S)$ which as $\alpha \cdot s \in \alpha \cdot S$ contradicts the fact that $\sup(\alpha \cdot S)$ is the supremum of $\alpha \cdot S$ so we must have $\sup(\alpha \cdot S) = \alpha \cdot \sup(S)$. □

Theorem 9.46. *If* $S, T \subseteq \mathbb{R}_{\mathbb{C}}$ *have a supremum then* $S + T = \{s + t | s \in S \wedge t \in T\}$ *has a supremum with* $\sup(S + T) = \sup(S) + \sup(T)$

Proof. Let $r \in S + T$ then $\exists s \in S, \exists t \in T$ such that $r = s + t$. Using the fact that the supremum is a upper bound we have that $r = s + t \leqslant \sup(S) + \sup(T)$. Using the definition of the supremum we have then that

$$\sup(S + T) \leqslant \sup(S) + \sup(T) \tag{9.3}$$

Assume now that $\sup(S+T) < \sup(S) + \sup(T)$ then if we take $\varepsilon = \sup(S) + \sup(T) - \sup(S+T)$ we have that $0 < \varepsilon$. So $\sup(S) - \frac{\varepsilon}{2} < \sup(S)$ and $\sup(T) - \frac{\varepsilon}{2} < \sup(T)$ and by the definition of the supremum there exists a $s \in S$ and a $t \in S$ such that $\sup(S) - \frac{\varepsilon}{2} < s \leqslant \sup(S) + \frac{\varepsilon}{2}$ and $\sup(T) - \frac{\varepsilon}{2} < t \leqslant \sup(T)$. So $s + t > \sup(S) - \frac{\varepsilon}{2} + \sup(T) - \frac{\varepsilon}{2} = \sup(S) + \sup(T) - \varepsilon = \sup(S) + \sup(T) - \sup(S) - \sup(T) + \sup(S+T) = \sup(S+T)$ or $s + t > \sup(S+T)$ which conflict the fact that the supremum is a upper bound. Hence $\sup(S) + \sup(T) \leqslant \sup(S+T)$ which together with (9.3) proves that

$$\sup(S+T) = \sup(S) + \sup(T)$$

□

Corollary 9.47. *If $S \subseteq \mathbb{R}_{\mathbb{C}}$ has a supremum and $x \in \mathbb{R}_{\mathbb{C}}$ then $S + x = \{s + x | s \in S\}$ has a supremum with* $\sup(S+x) = \sup(S) + x$

Proof. Define $T = \{x\}$ then $\sup(T) = x$ exists and $S + T = \{s + t | s \in S \wedge t \in T\} = \{s = x | s \in S\} = S + x$. Hence using 9.46 we have that $\sup(S+x)$ exists and $\sup(S+x) = \sup(S) + \sup(T) = \sup(S) + x$ □

Theorem 9.48. *If $S, T \subseteq \mathbb{R}_{\mathbb{C}}$ have a infinum then $S + T = \{s + t | s \in S \wedge t \in T\}$ has a infinum with* $\inf(S+T) = \inf(S) + \inf(T)$

Proof. Let $r \in S + T$ then $\exists s \in S, \exists t \in T$ such that $r = s + t$. Using the fact that the infinum is a lower bound we have that $\inf(S) + \inf(T) \leqslant s + t = r$. Using the definition of the infinum we have then that

$$\inf(S) + \inf(T) \leqslant \inf(S+T) \tag{9.4}$$

Assume now that $\inf(S) + \inf(T) < \inf(S+T)$ then if we take $\varepsilon = \inf(S+T) - \inf(S) - \inf(T)$ we have that $0 < \varepsilon$. So $\inf(S) < \inf(S) + \frac{\varepsilon}{2}$ and $\inf(T) < \inf(T) + \frac{\varepsilon}{2}$ and by the definition of the infinum there exists a $s \in S$ and a $t \in S$ such that $\inf(S) \leqslant s < \inf(S) + \frac{\varepsilon}{2}$ and $\inf(T) \leqslant t < \inf(T) + \frac{\varepsilon}{2}$. So $s + t < \inf(S) + \inf(T) + \varepsilon = \inf(S+t) + \inf(T) + \inf(S+T) - \inf(S) - \inf(T) = \inf(S+T)$ or $s + t < \inf(S+T)$ which conflict the fact that the infinum is a lower bound. Hence $\inf(S+T) \leqslant \inf(S) + \inf(T)$ which together with (9.4) proves that

$$\inf(S+T) = \inf(S) + \inf(T)$$

□

Corollary 9.49. *If $S \subseteq \mathbb{R}_{\mathbb{C}}$ has a infinum and $x \in \mathbb{R}_{\mathbb{C}}$ then $S + x = \{s + x | s \in S\}$ has a infinum with* $\inf(S+x) = \inf(S) + x$

Proof. Define $T = \{x\}$ then $\inf(T) = x$ exists and $S + T = \{s + t | s \in S \wedge t \in T\} = \{s = x | s \in S\} = S + x$. Hence using 9.48 we have that $\inf(S+x)$ exists and $\inf(S+x) = \inf(S) + \inf(T) = \inf(S) + x$ □

Theorem 9.50. $\langle \mathbb{Q}_{\mathbb{C}}, \leqslant \rangle$ *is not conditionally complete*

Proof. We prove this by contradiction, so assume that $\mathbb{Q}_{\mathbb{C}}$ is conditionally complete then as $\mathbb{Q}$ is order isomorphic with $\mathbb{Q}$ by 9.37 we have by 2.178 that $\mathbb{Q}$ is conditionally complete contradiction the fact that Q is not conditionally complete by 7.38. □

Corollary 9.51. *$\mathbb{Q}_{\mathbb{C}} \subset \mathbb{R}_{\mathbb{C}}$ [there exists a $r \in \mathbb{R}_{\mathbb{C}}$ such that $r \notin \mathbb{Q}_{\mathbb{C}}$]. In other words we have that $\mathbb{R}_{\mathbb{C}} \backslash \mathbb{Q}_{\mathbb{C}} \neq \emptyset$. $\mathbb{R}_{\mathbb{C}} \backslash \mathbb{Q}_{\mathbb{C}}$ is called the set of irrational numbers.*

Proof. As $\langle \mathbb{Q}_{\mathbb{C}}, \leqslant \rangle$ is not conditional complete (see previous theorem) there exists a non-empty set $A \subseteq \mathbb{Q}_{\mathbb{C}}$ with a upper bound u so that $\sup(A)$ does not exists in $\mathbb{Q}_{\mathbb{C}}$. Because $\mathbb{Q}_{\mathbb{C}} \subseteq \mathbb{R}_{\mathbb{C}}$ we have $\emptyset \neq A \subseteq \mathbb{R}_{\mathbb{C}}$ and that A has the upper bound $u \in \mathbb{Q}_{\mathbb{C}} \subseteq \mathbb{R}$. As $\mathbb{R}_{\mathbb{C}}$ is conditional complete (see 9.43) there exists a lowest upper bound $s = \sup(A)$, now if $s \in \mathbb{Q}_{\mathbb{R}}$ it would be a upper bound of A and if $b \in \mathbb{Q}_{\mathbb{R}}$ of A is another upper bound of A it is also a upper bound of A in $\mathbb{R}_{\mathbb{C}}$ and thus $s \leqslant b$ so s would be the supremum of A in $\mathbb{Q}_{\mathbb{R}}$ contradicting the fact that $\sup(A)$ does not exists in $\mathbb{Q}_{\mathbb{R}}$. So $s \notin \mathbb{Q}_{\mathbb{R}}$. □

Theorem 9.52. *We have the following properties concerning $\mathbb{N}_{0_{\mathbb{C}}}$ and $\mathbb{Z}_{\mathbb{C}}$ (here $-\mathbb{N}_{0_{\mathbb{C}}} = \{n \vdash -n \in \mathbb{N}_{0_{\mathbb{R}}}\} \subseteq \mathbb{Z}_{\mathbb{R}}$)*

1. *$\mathbb{Z}_{\mathbb{C}} = \mathbb{N}_{0_{\mathbb{C}}} \bigcup (-\mathbb{N}_{0_{\mathbb{C}}})$ and $\mathbb{N}_{0_{\mathbb{C}}} \bigcap (-\mathbb{N}_{0_{\mathbb{C}}}) = \{0\}$*
2. *If $n \in \mathbb{Z}_{\mathbb{C}}$ and $0 \leqslant n \Rightarrow n \in \mathbb{N}_{0_{\mathbb{C}}}$ (so $\mathbb{N}_{0_{\mathbb{C}}}$ is indeed the set of positive integers)*
3. *$\langle \mathbb{N}_{0_{\mathbb{C}}}, + \rangle$ is a semi-group of $\langle \mathbb{Z}_{\mathbb{C}}, + \rangle$*
4. *$\langle \mathbb{N}_{0_{\mathbb{C}}}, \cdot \rangle$ is a sub-semi-group of $\langle \mathbb{Z}_{\mathbb{C}}, \cdot \rangle$*

Proof.

1. As $\mathbb{N}_{0_{\mathbb{C}}}, -\mathbb{N}_{0_{\mathbb{C}}} \subseteq \mathbb{Z}_{\mathbb{C}}$ we have $\mathbb{N}_{0_{\mathbb{C}}} \bigcup (-\mathbb{N}_{0_{\mathbb{C}}}) \subseteq \mathbb{Z}_{\mathbb{C}}$. If $x \in \mathbb{Z}_{\mathbb{C}}$ then $\exists x' \in \mathbb{Z}_{\mathbb{R}}$ such that $x = (x', 0)$ as $\mathbb{Z}_{\mathbb{R}} \underset{8.58}{=} \mathbb{N}_{0_{\mathbb{R}}} \bigcup (-\mathbb{N}_{0_{\mathbb{R}}})$ we have that either $x' \in \mathbb{N}_{0_{\mathbb{R}}}$ or $-x' \in \mathbb{N}_{0_{\mathbb{R}}}$ hence $(x', 0) \in \mathbb{N}_{0_{\mathbb{C}}}$ or $(-x', 0) \in \mathbb{N}_{0_{\mathbb{C}}}$ proving that $x \in \mathbb{N}_{0_{\mathbb{C}}} \bigcup (-\mathbb{N}_{0_{\mathbb{C}}})$
2. If $n = (n', 0) \in \mathbb{Z}_{\mathbb{C}}$ and $0 \leqslant_{\mathbb{R}} n = (n', 0)$ then $0 \leqslant n' \underset{8.58}{=} n' \in \mathbb{N}_{0_{\mathbb{R}}} \Rightarrow n = (n', 0) \in \mathbb{N}_{0_{\mathbb{C}}}$
3. If $x, y \in \mathbb{N}_{0_{\mathbb{C}}}$ then we have $x = (x', 0), y = (y', 0)$ and $x', y' \in \mathbb{N}_{0_{\mathbb{R}}}$ so that
   a. $x + y = (x', 0) + (y', 0) = (x' + y', 0) \in \mathbb{N}_{0_{\mathbb{C}}}$ as $x' + y' \in \mathbb{N}_{0_{\mathbb{R}}}$ (for $\mathbb{N}_{0_{\mathbb{R}}}$ is a semi-group)
   b. $0 = (0, 0) \in \mathbb{N}_{0_{\mathbb{C}}}$
4. If $x, y \in \mathbb{N}_{0_{\mathbb{C}}}$ then we have $x = (x', 0), y = (y', 0)$ and $x', y' \in \mathbb{N}_{0_{\mathbb{R}}}$ so that
   a. $x \cdot y = (x' \cdot y' - 0 \cdot 0, x' \cdot 0 + 0 \cdot y') = (x' \cdot y', 0) \in \mathbb{N}_{0_{\mathbb{C}}}$ (as $x' \cdot y' \in \mathbb{N}_{0_{\mathbb{C}}}$ which is a semi-group)
   b. $1 = (1, 0) \in \mathbb{N}_{0_{\mathbb{C}}}$ □

Lemma 9.53. *If $n \in \mathbb{N}_{0_{\mathbb{C}}}$ and $0 < n \Rightarrow 1 \leqslant_{\mathbb{R}} n$.*

Proof. As $n = (n', 0)$ with $n' \in \mathbb{N}_{0_{\mathbb{R}}}$ we have from $0 \leqslant_{\mathbb{R}} n$ that $0 < n'$ hence using 8.60 we have $1 \leqslant n'$ proving that $1 = (1, 0) \leqslant (n', 0) = n$ □

Theorem 9.54. (Archimedean Property) *If $x \in \mathbb{R}_{\mathbb{C}}$ with $0 <_{\mathbb{R}} x$ then $\exists n \in \mathbb{N}_{0_{\mathbb{R}}}$ such that $y <_{\mathbb{R}} n \cdot x$*

Proof. As $x, y \in \mathbb{R}_{\mathbb{C}}$ there exists $x', y' \in \mathbb{R}$ such that $x = (x', 0)$, $y = (y', 0)$, using the Archimedean property of $\mathbb{R}$ (see 8.61) there exists a $n' \in \mathbb{N}_{0_{\mathbb{R}}}$ such that $y' < n' \cdot x'$ hence $y = (y', 0) <_{\mathbb{R}} (n' \cdot x', 0) = (n', 0) \cdot (x', 0) = n \cdot x$ where $n = (n', 0) \in \mathbb{N}_{0_{\mathbb{C}}}$ □

Theorem 9.55. *Let $x \in \mathbb{R}_{\mathbb{C}}$ then the following holds*

1. $\exists n \in \mathbb{N}_{0_{\mathbb{C}}} \vdash x <_{\mathbb{R}} n$
2. $\exists n \in \mathbb{Z}_{\mathbb{C}} \vdash n \leqslant_{\mathbb{R}} x <_{\mathbb{R}} n + 1$
3. *If* $0 <_{\mathbb{R}} x$ *then* $\exists n \in \mathbb{N}_{0_{\mathbb{R}}} \setminus \{0\}$ *such that* $\frac{1}{n} <_{\mathbb{R}} x$
4. *If* $0 \leqslant_{\mathbb{R}} x$ *then* $\exists n \in \mathbb{N}_{0_{\mathbb{R}}} \setminus \{0\}$ *such that* $n - 1 \leqslant_{\mathbb{R}} x <_{\mathbb{R}} n$

Proof. As $x \in \mathbb{R}_{\mathbb{C}}$ we have that $x = (x', 0)$ where $x' \in \mathbb{R}$

1. As $0 <_{\mathbb{R}} 1 \in \mathbb{R}_{\mathbb{C}}$ we have by the Archimedean property (see 9.54) that $\exists n \in \mathbb{N}_{0_{\mathbb{C}}}$ such that $x <_{\mathbb{R}} 1 \cdot n = n$
2. There exists by 8.62 (2) a $n' \in \mathbb{Z}_{\mathbb{R}}$ such that $n' \leqslant x' < n' + 1$ hence we have $(n', 0) \leqslant_{\mathbb{R}} (x', 0) <_{\mathbb{R}} (n' + 1, 0) = (n', 0) + (1, 0)$ or if we take $n = (n', 0) \in \mathbb{Z}_{\mathbb{C}}$ that $n \leqslant_{\mathbb{R}} x <_{\mathbb{R}} n + 1$
3. As $0 <_{\mathbb{R}} x = (x', 0)$ we must have $0 < x'$ and thus using 8.62 (3) there exists a $n' \in \mathbb{N}_{0_{\mathbb{R}}} \setminus \{0\}$ such that $\frac{1}{n'} < x'$ hence $\left(\frac{1}{n'}, 0\right) <_{\mathbb{R}} (x', 0) = x$ if we take $n = (n', 0) \in \mathbb{N}_{0_{\mathbb{C}}} \setminus \{0\}$ then using 9.2 we have that $\frac{1}{n} = n^{-1} = (n'^{-1}, 0) = \left(\frac{1}{n'}, 0\right)$ proving that $\frac{1}{n} <_{\mathbb{R}} x$
4. As $0 \leqslant_{\mathbb{R}} x$ we have $0 \leqslant x'$ we have by 8.62 that there exists $n' \in \mathbb{N}_{0_{\mathbb{R}}} \setminus \{0\}$ so that $n' - a \leqslant x' < n'$ hence we have $(n' - 1, 0) \leqslant (x', 0) \leqslant (n', 0)$ so $n - 1 \leqslant_{\mathbb{R}} x <_{\mathbb{R}} n$ if we define $n = (n', 0) \in \mathbb{N}_{0_{\mathbb{C}}}$ □

Corollary 9.56. *Let $x, y \in \mathbb{R}_{\mathbb{C}}$ then we have*

1. *If* $\forall \varepsilon \in \mathbb{R}_{\mathbb{C}_+}$ *we have* $x \leqslant_{\mathbb{R}} y + \varepsilon$ *then* $x \leqslant_{\mathbb{R}} y$
2. *If* $\forall \varepsilon \in \mathbb{R}_{\mathbb{C}_+}$ *we have* $x <_{\mathbb{R}} y + \varepsilon$ *then* $x \leqslant_{\mathbb{R}} y$
3. *If* $\forall n \in \mathbb{N}_{0_{\mathbb{C}}} \setminus \{0\}$ *we have* $x \leqslant_{\mathbb{R}} y + \frac{1}{n}$ *then* $x \leqslant_{\mathbb{R}} y$
4. *If* $\forall n \in \mathbb{N}_{0_{\mathbb{C}}} \setminus \{0\}$ *we have* $x <_{\mathbb{R}} y + \frac{1}{n}$ *then* $x \leqslant_{\mathbb{R}} y$
5. *If* $\forall \varepsilon \in \mathbb{R}_{\mathbb{C}_+}$, $a \in \mathbb{R}_{\mathbb{C}} \vdash 0 \leqslant a$ *we have* $x \leqslant_{\mathbb{R}} y + \varepsilon \cdot a$ *then* $x \leqslant_{\mathbb{R}} y$
6. *If* $\forall \varepsilon \in \mathbb{R}_{\mathbb{C}_+}$, $a \in \mathbb{R}_{\mathbb{C}} \vdash 0 \leqslant_{\mathbb{R}} a$ *we have* $x <_{\mathbb{R}} y + \varepsilon \cdot a$ *then* $x \leqslant_{\mathbb{R}} y$
7. *If* $\forall n \in \mathbb{N}_{0_{\mathbb{C}}} \setminus \{0\}$, $a \in \mathbb{R}_{\mathbb{C}} \vdash 0 \leqslant_{\mathbb{R}} a$ *we have* $x \leqslant_{\mathbb{R}} y + \frac{1}{n} \cdot a$ *then* $x \leqslant_{\mathbb{R}} y$
8. *If* $\forall n \in \mathbb{N}_{0_{\mathbb{C}}} \setminus \{0\}$, $a \in \mathbb{R}_{\mathbb{C}} \vdash 0 \leqslant_{\mathbb{R}} a$ *we have* $x <_{\mathbb{R}} y + \frac{1}{n} \cdot a$ *then* $x \leqslant y$

Proof.

1. Assume that $y <_{\mathbb{R}} x \Rightarrow 0 <_{\mathbb{R}} x - y$ then by the above theorem (3) there exists a $n \in \mathbb{N}_{0_{\mathbb{C}}} \setminus \{0\}$ such that for $\varepsilon = \frac{1}{n}$ we have $\varepsilon <_{\mathbb{R}} x - y \Rightarrow y + \varepsilon <_{\mathbb{R}} x \leqslant_{\mathbb{R}} y + \varepsilon \Rightarrow \varepsilon <_{\mathbb{R}} \varepsilon$ a contradiction so we must have $x \leqslant_{\mathbb{R}} y$

2. $x <_{\mathbb{R}} y + \varepsilon \Rightarrow x \leqslant_{\mathbb{R}} y + \varepsilon \underset{(1)}{\Rightarrow} x \leqslant_{\mathbb{R}} y$
3. If $\varepsilon \in \mathbb{N}_{0_{\mathbb{C}}} \backslash \{0\}$ then by the previous theorem there exists a $0 <_{\mathbb{R}} \frac{1}{n} <_{\mathbb{R}} \varepsilon \Rightarrow x \leqslant_{\mathbb{R}} y + \frac{1}{n} <_{\mathbb{R}} y + \varepsilon \underset{(2)}{\Rightarrow} x \leqslant_{\mathbb{R}} y$
4. $x <_{\mathbb{R}} y + \frac{1}{n} \Rightarrow x \leqslant_{\mathbb{R}} y + \frac{1}{n} \underset{(4)}{\Rightarrow} x \leqslant_{\mathbb{R}} y$
5. Given $\varepsilon \in \mathbb{R}_{\mathbb{C}_+}$ take $\varepsilon' = \frac{\varepsilon}{a+1} >_{\mathbb{R}} 0$ then from $x \leqslant_{\mathbb{R}} y + \varepsilon' \cdot a = y + \frac{a}{a+1} \cdot \varepsilon <_{\mathbb{R}} y + \varepsilon \underset{(2)}{\Rightarrow} x \leqslant_{\mathbb{R}} y$
6. $x <_{\mathbb{R}} y + \varepsilon \cdot a \Rightarrow x \leqslant_{\mathbb{R}} y + \varepsilon \cdot a \underset{(5)}{\Rightarrow} x \leqslant_{\mathbb{R}} y$
7. Given $n \in \mathbb{N}_{0_{\mathbb{C}}} \backslash (0)$ then as $0 <_{\mathbb{R}} \frac{1}{n \cdot (a+1)}$ there exists by the above theorem a $m \in \mathbb{N}_{0_{\mathbb{C}}} \backslash \{0\}$ such that $\frac{1}{m} <_{\mathbb{R}} \frac{1}{n \cdot (a+1)} \Rightarrow \frac{1}{m} \cdot a <_{\mathbb{R}} \frac{1}{n} \cdot \frac{a}{a+1} <_{\mathbb{R}} \frac{1}{n}$ so that $x \leqslant_{\mathbb{R}} y + \frac{1}{m} \cdot a <_{\mathbb{R}} y + \frac{1}{n} \underset{(3)}{\Rightarrow} x \leqslant_{\mathbb{R}} y$
8. $x <_{\mathbb{R}} y + \frac{1}{n} \cdot a \Rightarrow x \leqslant_{\mathbb{R}} y + \frac{1}{n} \cdot a \underset{(7)}{\Rightarrow} x \leqslant_{\mathbb{R}} y$ □

Theorem 9.57. (Density Theorem) *If $x, y \in \mathbb{R}_{\mathbb{C}}$ such that $x < y$ then $\exists q \in \mathbb{Q}_{\mathbb{C}}$ and $\exists r \in \mathbb{R}_{\mathbb{C}} \backslash \mathbb{Q}_{\mathbb{C}}$ such that $x <_{\mathbb{R}} q <_{\mathbb{R}} y$ and $x <_{\mathbb{R}} r <_{\mathbb{R}} y$*

Proof. First as $x, y \in \mathbb{R}_{\mathbb{C}}$ with $x < y$ there exists $x', y' \in \mathbb{R}$ such that $x = (x', 0)$, $y = (y', 0)$ and $x' < y'$. Using then the density of $\mathbb{Q}$ in $\mathbb{R}$ (see 8.64) there exists a $q' \in \mathbb{Q}_{\mathbb{C}}$ and $r' \in \mathbb{R} \backslash \mathbb{Q}_{\mathbb{C}}$ such that $x' < q' < y'$ and $x' < r' < y'$ and thus $(x', 0) <_{\mathbb{R}} (q', 0) <_{\mathbb{R}} (y', 0)$ and $(x', 0) <_{\mathbb{R}} (r', 0) <_{\mathbb{R}} (y', 0)$. Hence if we define $q = (q', 0) \in \mathbb{R}_{\mathbb{C}}$ and $r = (r', 0) \in \mathbb{R}_{\mathbb{C}} \backslash \mathbb{Q}_{\mathbb{C}}$ then we have $x <_{\mathbb{R}} q <_{\mathbb{R}} y$ and $x <_{\mathbb{R}} r <_{\mathbb{R}} y$ □

To simplify notation as we mix elements of $\mathbb{N}_0$ and $\mathbb{C}$ we introduce the following definition and notation.

Definition 9.58. *Let $n \in \mathbb{N}_0$ and $x \in \mathbb{R}_{\mathbb{C}}$ then $n \cdot x = \begin{cases} 0 \text{ if } n=0 \\ x + (n-1) \cdot x \text{ if } n > 1 \end{cases}$. Hence $0.x = 0$ (last 0 is the neutral element in $\mathbb{R}_{\mathbb{C}}$), $1 \cdot x = x + 0 \cdot x = x + 0 = x$, $2 \cdot x = x + 1 \cdot x = x + x$, ...*

The above operation is consistent with the multiplication in $\mathbb{R}_{\mathbb{C}}$ as the following theorem proves.

Theorem 9.59. *Let $n \in \mathbb{N}_0$ then $n \cdot x = i_{\mathbb{R}_{\mathbb{N}_0}}(n) \cdot x$ hence if 1 is the multiplicative unit in $\mathbb{R}_{\mathbb{C}}$ we have $n \cdot 1 = i_{\mathbb{R}_{\mathbb{N}_0}}(n)$*

Proof. By 9.35 we have that $i_{\mathbb{R}_{\mathbb{N}_0}}(0) = 0 \in \mathbb{R}_{\mathbb{C}}$ and $i_{\mathbb{Q}_{\mathbb{N}_0}}(1) = 1$. We prove now the theorem by induction so let $\mathcal{S} = \{n \in \mathbb{N}_0 | n \cdot x = i_{\mathbb{Q}_{\mathbb{N}_0}}(n) \cdot x\}$ then we have

$0 \in \mathcal{S}$. As $0 \cdot x = 0 = i_{\mathbb{R}_{\mathbb{N}_0}}(0) = i_{\mathbb{R}_{\mathbb{N}_0}}(0) \cdot x$

$n \in \mathcal{S} =< n+1 \in \mathcal{S}$. we have that $(n+1) \cdot x \underset{n+1>0}{=} x + ((n+1)-1) \cdot x \underset{n \in \mathcal{S}}{=} x + i_{\mathbb{R}_{\mathbb{N}_0}}(n) \cdot x = i_{\mathbb{R}_{\mathbb{N}_0}}(1) \cdot x + i_{\mathbb{R}_{\mathbb{N}_0}}(n) \cdot x = i_{\mathbb{R}_{\mathbb{N}_0}}(n+1) \cdot x = x$ □

Motivated by the above theorem and the fact that we already use the sames symbols for the neutral element and multiplicative unit in $\mathbb{R}$ and $\mathbb{N}_0$ we introduce the following notation.

Notation 9.60. *If $n \in \mathbb{N}_0$ then we note $i_{\mathbb{R}_{\mathbb{N}_0}}(n)$ also by n. So if in a operation (sum, multiplication or comparison) we mix elements of $\mathbb{N}_0$ and elements of $\mathbb{R}$ we implicitly assume that we have applied $i_{\mathbb{Q}_{\mathbb{N}_0}}$ to the elements of $\mathbb{N}_0$.*

Theorem 9.61. *$\forall n \in \mathbb{N}_{0_\mathbb{C}}$ we have that $n <_\mathbb{R} 2^n$ [and thus by 9.55 we have also $\forall a \in \mathbb{R}_\mathbb{C}\ \exists n \in \mathbb{N}_{0_\mathbb{C}}$ such that $a <_\mathbb{R} n_\mathbb{R} < 2^n$] (where $2 = (2,0) \in \mathbb{N}_{0_\mathbb{C}}$)*

Proof. This is proved by induction so let $S = \{n \in \mathbb{N}_0 | n <_\mathbb{R} 2^n\}$ then we have that

1. If $n = 0$ then $0 < 1 = 2^0 \Rightarrow 0 = 0 <_\mathbb{R} 1 = 2^0$ so that $0 \in S$
2. If $n \in S$ then $n <_\mathbb{R} 2^n$ then consider the following cases for $n+1$ (as $0 <_\mathbb{R} n + 1$)
   a. $\boldsymbol{(n+1=1)}$ then $n+1 = 1 <_\mathbb{R} 2 = 2^1 = 2^{n+1} \Rightarrow n+1 \in S$
   b. $\boldsymbol{(1 <_\mathbb{R} n+1)}$ then $2 \leqslant_\mathbb{R} n+1 \Rightarrow 1 \leqslant_\mathbb{R} n$ so that $n+1 \leqslant_\mathbb{R} n+n = 2 \cdot n <_\mathbb{R} 2 \cdot 2^n = 2^{n+1}$ proving that $n+1 \in S$ □

Theorem 9.62. *Let $x \in \mathbb{R}_\mathbb{C}$ with $x >_\mathbb{R} 1$ then $x^n - 1 \geqslant_\mathbb{R} n \cdot (x-1)\ \forall n \in \mathbb{N}_0$*

Proof. We prove this by induction, so let $S = \{n \in \mathbb{N}_0 | x^n - 1 \geqslant n \cdot (x-1)\}$ then we have

1. if $n = 0$ then $x^n - 1 = x^0 - 1 = 1 - 1 = 0 \geqslant_\mathbb{R} 0 = 0 \cdot (x-1) = n \cdot (x-1) \Rightarrow 0 \in S$
2. If $n \in S$ then $x^{n+1} - 1 = x \cdot x^{n+1} - 1 = x \cdot (x^n - 1) + (x-1) \geqslant_\mathbb{R} x \cdot (n \cdot (x-1)) + (x-1) = (x-1) \cdot (x \cdot n + 1) \geqslant_{\mathbb{R}_{x >_\mathbb{R} 1 \Rightarrow n \cdot x \geqslant_\mathbb{R} n}} (x-1) \cdot (n+1) \Rightarrow n+1 \in S$ □

Theorem 9.63. *If $N \in \mathbb{N}_{0_\mathbb{C}} \setminus \{0\}$ and $x \in \mathbb{R}_\mathbb{C}$ with $x >_\mathbb{R} 1$ then there exists a $n \in \mathbb{N}_{0_\mathbb{C}} \setminus \{0\}$ such that $x^n >_\mathbb{R} N$*

Proof. Consider the following cases for $N \in \mathbb{N}_{0_\mathbb{C}} \setminus \{0\}$

1. $\boldsymbol{(N=1)}$ then $N = 1 <_\mathbb{R} x = x^1$ proving the theorem
2. $\boldsymbol{(N>1)}$ take $\delta = \frac{N-1}{x-1} >_\mathbb{R} 0$ then by the Archimedean property (9.55) there exists a $n \in \mathbb{N}_{0_\mathbb{C}}$ such that $\delta <_\mathbb{R} n$ as by definition of δ we have $N - 1 = \delta \cdot (x-1) \leqslant_\mathbb{R} n \cdot (x-1) <_\mathbb{R} x^n - 1$ (using the previous theorem) so that $N <_\mathbb{R} x^n$ □

Corollary 9.64. *Let $x, y \in \mathbb{R}_\mathbb{C}$ with $x >_\mathbb{R} 1$ then there exists a $n \in \mathbb{N}_{0_\mathbb{C}} \setminus \{0\}$ such that $x^n >_R y$*

Proof. Using 9.55 there exists a $N \in \mathbb{N}_{0_\mathbb{C}} \setminus \{0\}$ such that $y <_\mathbb{R} N$, using the previous theorem there exists then a $n \in \mathbb{N}_{0_\mathbb{C}} \setminus \{0\}$ such that $N <_\mathbb{R} x^n$ hence $y <_\mathbb{R} x^n$ □

Theorem 9.65. *If $x \in \mathbb{R}_\mathbb{C}$ with $0 <_\mathbb{R} x <_\mathbb{R} 1$ and $n, m \in \mathbb{N}_{0_\mathbb{C}}$ with $n <_\mathbb{R} m$ then $x^m <_\mathbb{R} x^n$*

Proof. We prove this by induction, so let $m\in\mathbb{N}_{0_{\mathbb{C}}}$ and take $S_m=\{n\in\mathbb{N}|x^{m+n}<_{\mathbb{R}}x^m\}$ then we have

1. if $n=1$ then $x^{m+n}=x^{n+1}=x\cdot(x^n)<_{\mathbb{R}_{x<_{\mathbb{R}}1\Rightarrow x\cdot x^n<_{\mathbb{R}}x^n}}x^n\Rightarrow 1\in S_m$
2. if $n\in S_m$ then $x^{m+(n+1)}=x\cdot x^{m+n}<_{\mathbb{R}_{x_{\mathbb{R}}<1}}x^{m+n}<_{\mathbb{R}_{n\in S_m}}x^m$

Using mathematical induction we have then that $S_M=\mathbb{N}$. If now $n,m\in\mathbb{N}_{0_{\mathbb{R}}}$ with $n<_{\mathbb{R}}m$ then $k=m-n>_{\mathbb{R}}0$ or $k\in\mathbb{N}=S_n$ so that $x^m=x^{n+k}<_{\mathbb{R}}x^n$ $\square$

Theorem 9.66. *If* $\varepsilon\in\mathbb{R}_{\mathbb{C}_+}$ *and* $x\in\mathbb{R}_{\mathbb{C}}$ *such that* $0<_{\mathbb{R}}x<_{\mathbb{R}}1$ *then* $\exists N\in\mathbb{N}_{0_{\mathbb{C}}}\backslash\{0\}$ *such that* $0<_{\mathbb{R}}x^n<_{\mathbb{R}}\varepsilon$ *if* $n\geqslant N$

Proof. As $\varepsilon>_{\mathbb{R}}0$ we have that $\frac{1}{\varepsilon}$ is defined and by the Archimedean property (see 9.55) there exists a $n\in\mathbb{N}_{0_{\mathbb{C}}}$ such that $0<_{\mathbb{R}}\frac{1}{\varepsilon}<_{\mathbb{R}}n$, as $0<_{\mathbb{R}}x<_{\mathbb{R}}1$ we have that $1<_{\mathbb{R}}\frac{1}{x}$ so that by 9.63 there exists a $N\in\mathbb{N}_{0_{\mathbb{C}}}\backslash\{0\}$ such that $n<_{\mathbb{R}}x^N$ and thus we have that $0<_{\mathbb{R}}\frac{1}{\varepsilon}<_{\mathbb{R}}\left(\frac{1}{x}\right)^N=\frac{1}{x^N}\underset{x,\varepsilon>0}{\Rightarrow}0<_{\mathbb{R}}x^N<_{\mathbb{R}}\varepsilon$. If now we have $n\geqslant N$ then by the previous theorem we have $0<_{\mathbb{R}}x^n\leqslant_{\mathbb{R}}x^N<_{\mathbb{R}}\varepsilon$ $\square$

Theorem 9.67. *Let* $x,y\in\mathbb{R}_{\mathbb{C}+}$ *such that* $y<x$ *then* $\forall n\in\mathbb{N}$ *we have* $y^n<x^n$

Proof. We prove this by induction, so take $\mathcal{S}=\{n\in\mathbb{N}|0<y^n<x^n\}$ then we have

$\mathbf{1}\in\boldsymbol{\mathcal{S}}$**.** then $0<y^1=y<x=x^1$ proving that $1\in\mathcal{S}$

$\boldsymbol{n}\in\boldsymbol{\mathcal{S}}\Rightarrow\boldsymbol{n}+\mathbf{1}\in\boldsymbol{\mathcal{S}}$**.** then as $n\in\mathcal{S}$ we have $0<y^n<x^n$ then as $0<y$ we have that $0<y^n\cdot y<x^n\cdot y\Rightarrow 0<y^{n+1}<x^n\cdot y$, further as $y<x$ and $0<y^n<x^n$ we have $y\cdot x^n<x^n\cdot x=x^{n+1}$. Combining these two inequalities gives $0<y^{n+1}<x^{n+1}$ so that $n+1\in\mathcal{S}$. $\square$

Corollary 9.68. *Let* $x\in\mathbb{R}_{\mathbb{C}_+}$ *then* $\forall n\in\mathbb{N}_0$ *we have* $0<x^n$

Proof. If $n=0$ then $0<1=x^0$, if $n\in\mathbb{N}$ then using the previous theorem we have $0=0^n<x^0$ $\square$

9.4.3 Square root on $\mathbb{R}_{\mathbb{C}}$

Definition 9.69. *Let* $x=(x',0)\in\mathbb{R}_{\mathbb{C}}$ *with* $0\leqslant_{\mathbb{R}}x$ *then we define* $\sqrt{x}=\left(\sqrt{x'},0\right)$ *(which is well defined as from* $x\geqslant 0$ *it follows that* $x'\geqslant 0$*)*

Theorem 9.70. *Let* $^2\colon\{x\in\mathbb{R}_{\mathbb{C}}|0\leqslant_{\mathbb{R}}x\}\to\{x\in\mathbb{R}_{\mathbb{C}}|0\leqslant_{\mathbb{R}}x\}$ *be the map defined by* $x\to x^2$ *then* $\sqrt{\ }\colon\{x\in\mathbb{R}_{\mathbb{C}}|0\leqslant_{\mathbb{R}}x\}\to\{x\in\mathbb{R}_{\mathbb{C}}|0\leqslant_{\mathbb{R}}x\}$ *defined by* $x\to\sqrt{x}$ *is the inverse of* ^{2}*. Hence* ^{2} *and* $\sqrt{\ }$ *are bijections.*

Proof. Let $z=(z',0)\in\{x\in\mathbb{R}_{\mathbb{C}}|0\leqslant x\}$ then $(\sqrt{\ }\circ{}^2)(z)=\sqrt{(z^2)}=\sqrt{(z',0)\cdot(z',0)}=\sqrt{((z')^2,0)}=\left(\sqrt{(z')^2},0\right)=(z',0)=z$. Further we have $({}^2\circ\sqrt{\ })(z)=(\sqrt{z})^2=\left(\sqrt{z'},0\right)^2=\left(\sqrt{z'},0\right)\cdot\left(\sqrt{z'},0\right)=\left(\left(\sqrt{z'}\right)^2,0\right)=(z',0)=z$ $\square$

The restriction to positive real numbers in the above is needed because for example $(-1)^2=1=1^2$ and $i^2=-1=(-i)^2$.

Theorem 9.71. *$\sqrt{}:\{x\in\mathbb{R}|0\leqslant x\}\rightarrow\{x\in\mathbb{R}|0\leqslant x\}$ is a strictly increasing function*

Proof. If $x,y\in\{x\in\mathbb{R}|0\leqslant x\}$ is such that $x<y$. Assume then that $\sqrt{y}\leqslant\sqrt{x}$ then by 9.41 we have $(\sqrt{y})^2\leqslant(\sqrt{x})^2\Rightarrow y\leqslant x$ contradicting $x<y$ so we must have $\sqrt{x}<\sqrt{y}$ □

Theorem 9.72. *If $x,y\in\{x\in\mathbb{R}|0\leqslant x\}$ then $\sqrt{x\cdot y}=\sqrt{x}\cdot\sqrt{y}$*

Proof. As $(\sqrt{x\cdot y})^2=x\cdot y=(\sqrt{x})^2\cdot(\sqrt{y})^2=(\sqrt{x}\cdot\sqrt{y})^2$ we have by the fact that ^{2}: $\{x\in\mathbb{R}|0\leqslant x\}\rightarrow\{x\in\mathbb{R}|0\leqslant x\}$ is a bijection and thus injective so we have $\sqrt{x\cdot y}=\sqrt{x}\cdot\sqrt{y}$ □

Theorem 9.73. *Given $x,y\in\{x\in\mathbb{R}|0\leqslant x\}$ then we have $\sqrt{x+y}\leqslant\sqrt{x}+\sqrt{y}$*

Proof. We prove this by contradiction, so assume that $\sqrt{x}+\sqrt{y}<\sqrt{x+y}\underset{9.41}{\Rightarrow}-\sqrt{x+y}<-(\sqrt{x}+\sqrt{y})$ and by multiplying by $\sqrt{x}+\sqrt{y}\geqslant 0$ we have $-\sqrt{x+y}\cdot(\sqrt{x}+\sqrt{y})<-(\sqrt{x}+\sqrt{y})\cdot(\sqrt{x}+\sqrt{y})=-(x+y+2\cdot\sqrt{x}\cdot\sqrt{y})$. Now we have by 9.41 that $0\leqslant(\sqrt{x+y}-(\sqrt{x}+\sqrt{y}))^2=(\sqrt{x+y})^2+(\sqrt{x}+\sqrt{y})^2-2\cdot\sqrt{x+y}\cdot(\sqrt{x}+\sqrt{y})=x+y+(\sqrt{x})^2+(\sqrt{y})^2+2\cdot\sqrt{x}\cdot\sqrt{y}-2\cdot\sqrt{x+y}\cdot(\sqrt{x}+\sqrt{y})=2\cdot x+2\cdot y+2\cdot\sqrt{x}\cdot\sqrt{y}-2\cdot\sqrt{x+y}\cdot(\sqrt{x}+\sqrt{y})\underset{\text{multiply by } \frac{1}{2}}{\Rightarrow}0\leqslant x+y+\sqrt{x}\cdot\sqrt{y}-\sqrt{x+y}\cdot(\sqrt{x}+\sqrt{y})<x+y+\sqrt{x}\cdot\sqrt{y}-(x+y+2\cdot\sqrt{x}\,\sqrt{y})=-\sqrt{x}\cdot\sqrt{y}\leqslant 0\Rightarrow 0<0$ a contradiction. So we must have $\sqrt{x+y}\leqslant\sqrt{x}+\sqrt{y}$ □

We have now the biggest reason to introduce the complex numbers

Theorem 9.74. *If $z\in\mathbb{R}_{\mathbb{C}}$ then there exists a $x\in\mathbb{C}$ such that $z=x^2$*

Proof. We have two cases to consider for z

$0\leqslant z$. take then $x=\sqrt{z}$ hence $x^2=(\sqrt{z})^2=z$

$z<0$. then $0\leqslant(-1)\cdot z$ take then $x=i\cdot\sqrt{(-1)\cdot z}$ then $x^2=i^2\cdot\left(\sqrt{(-1)\cdot z}\right)^2=(-1)\cdot(-1)\cdot z=z$ □

9.5 Summarize

To summarize what we have done in the above:

1. $\langle\mathbb{C},+,\cdot\rangle$ forms a field
2. $\langle\mathbb{C},\leqslant_{\mathbb{C}}\rangle$ is a total ordered set
3. We have embedded $\mathbb{N}_0,\mathbb{Z},\mathbb{Q}$ and $\mathbb{R}$ in $\mathbb{C}$ as the sets $\mathbb{N}_{0_{\mathbb{C}}},\mathbb{Z}_{\mathbb{C}},\mathbb{Q}_{\mathbb{C}}$ and $\mathbb{R}_{\mathbb{C}}$ such that
    a. $\mathbb{N}_{0_{\mathbb{C}}}\subseteq\mathbb{Z}_{\mathbb{C}}\subseteq\mathbb{Q}_{\mathbb{C}}\subseteq\mathbb{C}$
    b. $\langle\mathbb{N}_{0_{\mathbb{C}}},+\rangle$ is a sub semi group of $\langle\mathbb{C},+\rangle$
    c. $\langle\mathbb{N}_{0_{\mathbb{C}}},\cdot\rangle$ is a sub semi group of $\langle\mathbb{C},\cdot\rangle$

d. $\langle \mathbb{Z}_{\mathbb{C}}, +, \cdot \rangle$ is a sub ring of $\langle \mathbb{C}, +, \cdot \rangle$

e. $\langle \mathbb{R}_{\mathbb{C}}, +, \cdot \rangle$ is a sub field of $\langle \mathbb{C}, +, \cdot \rangle$

f. $\langle \mathbb{R}_{\mathbb{C}}, \leqslant_R \rangle$ is a conditional complete total ordered set

g. $\mathbb{R}_{\mathbb{C}}$ has the Archimedean property

h. $\mathbb{Q}_{\mathbb{C}}$ is dense in $\mathbb{R}_{\mathbb{C}}$

i. $\forall z \in \mathbb{R}_{\mathbb{C}}$ there exists a $x \in \mathbb{R}$ such that $x^2 = z$ (later we will extend this by proving the fundamental theorem of Algebra)

j. $\mathbb{R}_{\mathbb{C}}^2$ is bijection with $\mathbb{C}$

4. We have defined a norm $||$ on $\mathbb{C}$ such that

   a. $0 \leqslant |z| \in \mathbb{R}_{\mathbb{C}}$

   b. $|z \cdot z'| = |z| \cdot |z'|$

   c. $\mathrm{Re}(z) \leqslant |z|$

   d. $z \cdot \bar{z} = |z|^2$

   e. $|z + y| \leqslant |x| + |y|$

   f. $|z| = 0 \Leftrightarrow z = 0$

   g. If $z \in \mathbb{R}_{\mathbb{C}}$ then $|z| = |\mathrm{Re}(z)|$

For the rest of this book we don't have to talk anymore about the original $\mathbb{Z}, \mathbb{Q}$, $\mathbb{R}$ (with the exception of $\mathbb{N}_0$ that we need for induction). So to simplify notation we will from now on use $\mathbb{N}_0, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ when we actually means $\mathbb{N}_{0_{\mathbb{C}}}, \mathbb{Q}_{\mathbb{C}}, \mathbb{R}_{\mathbb{C}}, \mathbb{C}$ and $\leqslant$ when we use $\leqslant_{\mathbb{R}}$ so that $\mathbb{N}_0 \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$.

Chapter 10

Linear Algebra

10.1 Sums and products

Definition 10.1. *If $n \in \mathbb{N}_0$ then $\{0, \ldots n\} = S_{n+1} = \{i \in \mathbb{N}_0 | i < n+1\} = n+1$ (see 5.16)*

if $1 \leqslant n$ then $\{1, \ldots n\} = \{0, \ldots, n\} \backslash \{0\}$ [if $n = 0 \Rightarrow \{1, \ldots, n\} = \emptyset$ and also if $n = 0$ then $\{1, \ldots, n-1\} = S_n = \emptyset$]

Definition 10.2. (Finite sum) *Let $\langle A, + \rangle$ be a semi-group, and $x = \{x_i\}_{i \in \{0, \ldots n\}}$ a family of elements of A (see 2.100) then $\forall j \in \{0, \ldots, n\}$ $\sum_{i=0}^{j} x_i$ is defined as follows via recursion (see 5.17) as follows. Define $g: S_n \times A \rightarrow A$ by $(i, a) \rightarrow a + x_{i+1}$ (note if $n = 0$ then g is the empty function) then we can use 5.17 to define a f: $S_{n+1} = \{0, \ldots, n\} \rightarrow A$ such that*

$$\begin{array}{rcl} f(0) & = & x_0 \\ \forall j \in S_n \text{ we have } f(j+1) & = & g(j, f(j)) \end{array}$$

we define then $\sum_{i=0}^{j} x_i = f(j)$.

Note 10.3. The index i used in $\sum_{i=0}^{j} x_i$ and $x = \{x_i\}_{i \in \{0, \ldots, n\}}$ is actually part of a bad (but traditional notation), we could easily use k instead of i. Probably a better notation would be $\sum_{\{0, \ldots, n\}} x$ (so no index is used). However the use of indexes allows use to use formula's based on the index, like in $x = \{10^i\}_{i \in \{1, \ldots, n\}}$ and $\sum_{i=0}^{n} 10^i$ so we will stick with this notation.

Note 10.4. If $\langle A, \cdot \rangle$ is a semi-group, where the operator $\cdot$ is interpreted as a product and not a sum, then we note $\sum_{i=n}^{m} x_i$ as $\prod_{i=n}^{m} x_i$

Theorem 10.5. *If $\langle A, + \rangle$ is a semi-group and $\{x_i\}_{i \in \{0, \ldots n\}}$ then we have*

1. $\sum_{i=0}^{0} x_i = x_0$
2. $\forall j \in \{1, \ldots, n\}$ *we have* $\sum_{i=0}^{j} x_i = (\sum_{i=0}^{j-1} x_i) + x_j$
3. $\forall j \in \{0, \ldots, n-1\}$ *we have* $\sum_{i=0}^{j+1} x_i = (\sum_{i=0}^{j} x_i) + x_{j+1}$

Proof.

1. $\sum_{i=0}^{0} x_i = f(0) = x_0$
2. *If* $j \in \{1, \ldots, n\}$ *then* $j = (j-1)+1$ *and* $\sum_{i=0}^{j} x_i = f((j-1)+1) = g(j-1, f(j-1)) = f(j-1)+x_j = \sum_{i=0}^{j-1} x_i + x_j$
3. *If* $j \in S_n$ *we have* $\sum_{i=0}^{j+1} x_i = f(j+1) = g(j, f(j)) = f(j)+x_{j+1} = (\sum_{i=0}^{j} x_i) + x_{j+1}$ □

Example 10.6. If $\langle A, +\rangle$ is a semi-group (additive with neutral element 0) and $\{x_i\}_{i\in\{0,\ldots,n\}}$ is such that $\forall i \in \{0, \ldots, n\}$ we have $x_i = a \in A$ then we have $\sum_{i=0}^{n} x_i = a \cdot (n+1)$ (see 4.22)

Proof. If $n \in \mathbb{N}_0$ then we have $X = \{n \in \mathbb{N}_0 |\text{if } \{x_i\}_{i\in\{0,\ldots,n\}} \text{ is such that } \forall i \in \{0, \ldots, n\} \text{ we have } x_i = a \text{ and thus } \sum_{i=0}^{n} x_i = a \cdot (n+1)\}$ then :

1. Then if $n = 0$ we have $\sum_{i=0}^{n} x_i = \sum_{i=0}^{0} x_i = x_0 = a = a \cdot 1 \Rightarrow 0 \in X$
2. If $n \in X$ then if $\{x_i\}_{i\in\{0,\ldots,n+1\}}$ is such that $\forall i \in \{0, \ldots, n+1\}$ we have $x_i = a$ then we have $\sum_{i=0}^{n+1} x_i = (\sum_{i=0}^{n} x_i) + x_{n+1} \underset{n\in X}{=} (a \cdot n) + a = a \cdot (n+1) \Rightarrow n+1 \in X$

Using mathematical induction we have then $X = \mathbb{N}_0$ proving the theorem. □

Example 10.7. If $\langle A, \cdot\rangle$ is a semi-group (multiplicative with neutral element 1) and $\{x_i\}_{i\in\{0,\ldots,n\}}$ is such that $\forall i \in \{0, \ldots, n\}$ we have $x_i = a \in A$ then we have $\prod_{i=0}^{n} x_i = a^{n+1}$ (see 4.22)

Proof. If $n \in \mathbb{N}_0$ then we have $X = \{n \in \mathbb{N}_0 |\text{if } \{x_i\}_{i\in\{0,\ldots,n\}} \text{ is such that } \forall i \in \{0, \ldots, n\} \text{ we have } x_i = a \text{ and thus } \prod_{i=0}^{n} x_i = a^{n+1}\}$ then :

1. Then if $n = 0$ we have $\prod_{i=0}^{n} x_i = \prod_{i=0}^{0} x_i = x_0 = a = a \cdot 1 \Rightarrow 0 \in X$
2. If $n \in X$ then if $\{x_i\}_{i\in\{0,\ldots,n+1\}}$ is such that $\forall i \in \{0, \ldots, n+1\}$ we have $x_i = a$ then we have $\prod_{i=0}^{n+1} = (\prod_{i=0}^{n} x_i) \cdot x_{n+1} \underset{n\in X}{=} (a^n) \cdot a = a^{n+1} \Rightarrow n+1 \in X$

Using mathematical induction we have then $X = \mathbb{N}_0$ proving the theorem. □

Lemma 10.8. *If* $\langle A, +\rangle$ *is a a semi-group,* $n \in \mathbb{N}_0$, $j \in \{0, \ldots, n\}$ *and* $\{x_i\}_{i\in\{0,\ldots,n\}}$ *a family in* A *(defined by* $x: \{0, \ldots, n\} \rightarrow A$ *so we can take* $x_{|\{0,\ldots,j\}}: \{0, \ldots, j\} \rightarrow A$*) then* $\sum_{k=0}^{n} x_k = \sum_{k=0}^{j} (x_{|\{0,\ldots,j\}})_k$ *. In the future we write* $\{(x_{|\{0,\ldots j\}})_i\}_{i\in\{0,\ldots,j\}}$ *as* $\{x_i\}_{i\in\{0,\ldots j\}}$ *and call it a sub family of* $\{x_i\}_{i\in\{0,\ldots,n\}}$

Proof. We prove this by induction on j so let $B = \{j \in \mathbb{N}_0 |\text{if } j \in \{0, \ldots n\} \text{ and } \{x_i\}_{i\in\{0,\ldots,n\}} \text{ a family in } A \text{ then } \sum_{k=0}^{j} x_k = \sum_{k=0}^{j} (x_{|\{0,\ldots,j\}})_k\}$ then we have

1. $\sum_{k=0}^{0} x_k = x_0 = x(0) = x_{|\{0,\ldots,j\}}(0) = (x_{|\{0,\ldots,j\}}) = \sum_{k=0}^{0} (x_{|\{0,\ldots,j\}})_k$ and thus $0 \in B$

2. If $j \in B$ then for $j+1$ we have

   a. **$(j+1 \notin \{0, \ldots, n\})$** then $j+1 \in B$

   b. **$(j+1 \in \{0, \ldots n\})$** then $\sum_{k=0}^{j+1} (x_{|\{0,\ldots,j+1\}})_k = (\sum_{k=0}^{j} (x_{|\{0,\ldots,j+1\}})_k) + (x_{|\{0,\ldots,j+1\}})_{j+1} \underset{j\in B}{=} (\sum_{k=0}^{j} ((x_{|\{0,\ldots,j+1\}})_{|\{0,\ldots,j\}})_k) + x_{j+1} \underset{(x_{|\{0,\ldots,j+1\}_{|\{0,\ldots,j\}}} = x_{|\{0,\ldots,,j\}})}{=} (\sum_{k=0}^{j} (x_{|\{0,\ldots,j\}})_k) + x_{j+1} \underset{j\in B}{=} (\sum_{k=0}^{j} x_j) + x_{j+1} = \sum_{k=0}^{j+1} x_k$ so $j+1 \in B$

Using mathematical induction (see 4.10) we have $B = \mathbb{N}_0$ and thus if $j \in \{0, \ldots, n\} \Rightarrow j \in \mathbb{N}_0 = B \Rightarrow \sum_{k=0}^{j} x_k = \sum_{k=0}^{j} (x_{|\{0,\ldots,j\}})_k$ □

Example 10.9. If $x = \{i\}_{i\in\{0,\ldots 3\}}$ then

$$\begin{aligned}
\sum_{i=0}^{3} i &= \left(\sum_{i=0}^{2} i\right) + 3\\
&= \left(\left(\sum_{i=0}^{1} i\right) + 2\right) + 3\\
&= \left(\left(\left(\sum_{i=0}^{0} i\right) + 1\right) + 2\right) + 3\\
&= ((0+1)+2)+3 = 6
\end{aligned}$$

Lemma 10.10. *If $\{\langle A_i, +_i\rangle\}_{i\in I}$ is a family of abelian groups then by 3.15 we have that $\langle \prod_{i\in I} A_i, +\rangle$ is a abelian group (here $(x+y)(i) \underset{\text{notation}}{=} (x+y)_i \underset{\text{definition}}{=} x_i +_i y_i$ and if $\{x_j\}_{j\in\{0,\ldots,n\}}$ is a family in $\prod_{i\in I} A_i$ then $\forall i \in I$ we have $(\sum_{i=0}^{n} x_j)(i) \underset{\text{notation}}{=} (\sum_{i=0}^{n} x_j)_i = \sum_{i=0}^{n} (x_j)(i) \underset{\text{notation}}{=} \sum_{i=0}^{n} (x_j)_i$ (last sum is based on $+_i$)*

Proof. This is easily proved by induction so let $S = \{n \in \mathbb{N}_0 | \text{if } \{x_j\}_{j\in\{0,\ldots,n\}} \text{ is a family in } \prod_{i\in I} A_i \text{ then } (\sum_{i=0}^{n} x_j)(i) = \sum_{i=0}^{n} (x_j)(i)\}$ then we have:

1. $(\sum_{j=0}^{0} x_j)(i) = (x_0)(i) = \sum_{j=0}^{0} (x_o)(i) \Rightarrow 0 \in S$

2. If $n \in S$ take then a family $\{x_i\}_{i\in\{0,\ldots,n+1\}}$ then we have $(\sum_{j=0}^{n+1} x_j)(i) = ((\sum_{j=0}^{n} x_j) + x_{n+1})(i) = (\sum_{j=0}^{n} x_j)(i) +_i x_{n+1}(i) \underset{n\in S}{=} (\sum_{j=0}^{n} x_j(i)) +_i x_{n+1}(i) = \sum_{j=0}^{n+1} x_j(i) \Rightarrow n+1 \in S$

So by mathematical induction we have $S = \mathbb{N}_0$ completing our proof. □

Lemma 10.11. *If $\langle A, +\rangle$ is a abelian semi-group and $\{x_i\}_{i\in\{0,\ldots,n\}}$, $\{y_i\}_{i\in\{0,\ldots,n\}}$ are families of elements in A then $\sum_{i=0}^{n} (x_i + y_i) = (\sum_{i=0}^{n} x_i) + (\sum_{i=0}^{n} y_i)$*

Proof. We prove this by induction on n so let $X = \{n \in \mathbb{N} | \text{if } \{x_i\}_{i \in \{0,\ldots,n\}}, \{y_i\}_{i \in \{0,\ldots,n\}} \text{ then } \sum_{i=0}^{n} (x_i + y_i) = (\sum_{i=0}^{n} x_i) + (\sum_{i=0}^{n} y_i)\}$ then we have:

1. If $n = 0$ then $\sum_{i=0}^{0} (x_i + y_i) = x_0 + y_0 = (\sum_{i=0}^{0} x_i) + (\sum_{i=0}^{0} y_i) \Rightarrow 0 \in X$
2. Assume that $n \in X$ then if $\{x_i\}_{i \in \{0,\ldots,n+1\}}, \{y_i\}_{i \in \{0,\ldots,n+1\}}$ we have $\sum_{i=0}^{n+1} (x_i + y_i) = (\sum_{i=0}^{n} (x_i + y_i)) + (x_{n+1} + y_{n+1}) \underset{n \in X}{=} ((\sum_{i=0}^{n} x_i) + (\sum_{i=0}^{n} y_i)) + (x_{n+1} + y_{n+1}) \underset{\text{commutativity,associativity}}{=} ((\sum_{i=0}^{n} x_i) + x_{n+1}) + ((\sum_{i=0}^{n} y_i) + y_{n+1}) = (\sum_{i=0}^{n+1} x_i) + (\sum_{i=0}^{n+1} y_i)$ proving that $n + 1 \in X$

and using mathematical induction we have $X = \mathbb{N}$ proving the theorem. $\square$

Theorem 10.12. *Let $\langle A, + \rangle$ be a abelian semi-group and $n, m \in \mathbb{N}_0$ and $\{x_{i,j}\}_{(i,j) \in \{0,\ldots,n\} \times \{0,\ldots,m\}}$ then we have $\sum_{i=0}^{n} (\sum_{j=0}^{m} x_{i,j}) = \sum_{j=0}^{m} (\sum_{i=0}^{n} x_{i,j})$*

Proof. We prove this by induction so let $S_n = \{m \in \mathbb{N}_0 | \text{if } \{x_{i,j}\}_{(i,j) \in \{0,\ldots,n\} \times \{0,\ldots,m\}} \text{ is } a \text{ family in } A \text{ then } \sum_{i=0}^{n} (\sum_{j=0}^{m} x_{i,j}) = \sum_{j=0}^{m} (\sum_{i=0}^{n} x_{i,j})\}$ then we have :

1. If $m = 0$ then $\sum_{i=0}^{n} (\sum_{j=0}^{0} x_{i,j}) = \sum_{i=0}^{n} x_{i,0} = \sum_{j=0}^{0} (\sum_{i=0}^{n} x_{i,j}) \Rightarrow 0 \in S_n$
2. Assume that $m \in S_n$ and take $\{x_{i,j}\}_{(i,j) \in \{0,\ldots,n\} \times \{0,\ldots,m+1\}}$ then $\sum_{i=0}^{n} (\sum_{j=0}^{m+1} x_{i,j}) = \sum_{i=0}^{n} ((\sum_{j=0}^{m} x_{i,j}) + x_{i,m+1}) \underset{\text{previous theorem}}{=} \sum_{i=0}^{n} (\sum_{j=0}^{m} x_{i,j}) + \sum_{i=0}^{n} x_{i,m+1} \underset{m \in S_n}{=} \sum_{j=0}^{m} (\sum_{i=0}^{n} x_{i,j}) + \sum_{i=0}^{n} x_{i,m+1} = \sum_{j=0}^{m+1} (\sum_{i=0}^{n} x_{i,j}) \Rightarrow m + 1 \in \mathcal{S}_n$

using mathematical induction we have that $S_n = \mathbb{N}_0$ proving the theorem. $\square$

Definition 10.13. *If I is a set and $\sigma: I \rightarrow I$ a bijective function then σ is called a* ***permutation***. *The set of all the permutations on I is noted as $S_I = \{f \in I^I | f \text{ is } a \text{ permutation}\}$*

Theorem 10.14. *Let I be a set and S_I its permutation set then $\langle S_I, \circ \rangle$ is a group*

Proof. First $\circ: S_I \times S_I \rightarrow S_I$ is a function as if $(f, g) \in S_I \times S_I$ we have that $f \circ g \in S_I$ (see 2.33)

1. if $f, g, h \in S_I$ then $(f \circ g) \circ h = f \circ (g \circ h)$ (see 2.34)
2. if $f \in S_I$ then $1_A \in S_I$ and by 2.34 we have $1_I \circ f = f = f \circ 1_I$
3. if $f \in S_I$ then as f is a bijection and we have thus that $f \circ f^{-1} = 1_A = f^{-1} \circ f$ $\square$

Theorem 10.15. *If $\sigma: I \rightarrow I$ is a permutation such that $\sigma(i) = i$ then $\sigma_{|I \setminus \{i\}}: I \setminus \{i\} \rightarrow I \setminus \{i\}$ is a permutation*

Proof.

1. **(injective)** If $\sigma_{|\mathrm{I}\backslash\{i\}}(j)=\sigma_{|I\backslash\{i\}}(j')\Rightarrow\sigma(j)=\sigma(j')\underset{\sigma\,\mathrm{is\,injective}}{\Rightarrow}j=j'$
2. **(surjective)** If $l\in I\backslash\{i\}\underset{\sigma\,\mathrm{is\,surjective}}{\Rightarrow}\exists k\in I\vdash\sigma(k)=l$ if $k=i$ then $l=\sigma(i)=i$ contradicting $l\in I\backslash\{i\}$ so $k\in I\backslash\{i\}$ and $\sigma_{|I\backslash\{i\}}(k)=\sigma(k)=l$ □

Theorem 10.16. *If $n\in\mathbb{N}_0$ and $i,j\in\{0,\ldots,n\}$ then $(i\leftrightarrow_n j)\colon\{0,\ldots,n\}\to\{0,\ldots,n\}$ defined by $(i\leftrightarrow_n j)(k)=\begin{cases}k\text{ if }k\neq i,j\\ j\text{ if }k=i\\ i\text{ if }k=j\end{cases}$ is a **permutation** called a **transposition**. Note that trivially $(i\leftrightarrow_n j)=(j\leftrightarrow_n i)$ and that $(i\leftrightarrow_n i)=i_{\{0,\ldots,n\}}$*

Proof. We have to prove injectivity and surjectivity

1. **(injectivity)** If $(i\leftrightarrow_n j)(k)=(i\leftrightarrow_n j)(k')$ then we have the following cases for k,k'
   a. $(\boldsymbol{k=i,k'=i})$ Then we have evidently that $k=k'$
   b. $(\boldsymbol{k=i,k'=j})$ Then $k'=j=(i\leftrightarrow_n j)(k)=(i\leftrightarrow_n j)(k')=i=k$ and thus $k=k'$
   c. $(\boldsymbol{k=i,k'\neq i,j})$ Then $j=(i\leftrightarrow_n j)(k)=(i\rightarrow_n j)(k')=k'\neq i,j$ a impossibility so this case does not count.
   d. $(\boldsymbol{k=j,k'=i})$ Then $k'=i=(i\leftrightarrow_n j)(k)=(i\leftrightarrow_n j)(k')=j=k$ and thus $k=k'$
   e. $(\boldsymbol{k=j,k'=j})$ Then we have evidently that $k=k'$
   f. $(\boldsymbol{k=j,k'\neq i,j})$ The $i=(i\leftrightarrow_n j)(k)=(i\leftrightarrow_n j)(k')=k'\neq i,j$ a impossibility so this case does not count.
   g. $(\boldsymbol{k\neq i,j,k'\neq i,j})$ Then $k=(i\rightarrow_n j)(k)=(i\rightarrow_n j)(k')=k'$ and thus $k=k'$.

   in all the cases that apply we have $k=k'$.
2. **(surjectivity)** If $l\in\{0,\ldots,n\}$ then we have the following cases for l
   a. $(\boldsymbol{l=i})$ then $l=(i\leftrightarrow_n j)(j)$
   b. $(\boldsymbol{l=j})$ then $l=(i\leftrightarrow_n j)(i)$
   c. $(\boldsymbol{l\neq i,j})$ then $l=(i\leftrightarrow_n j)(l)$

   proving surjectivity. □

Lemma 10.17. *If $n\in\mathbb{N}_0$ then if $(i,j)\in\{0,\ldots,n\}$ we have that $(i\leftrightarrow_{n+1}j)_{|\{0,\ldots,n\}}=(i\leftrightarrow_n j)$*

Proof. If $k\in\{0,\ldots,n\}$ then we have either

1. $(\boldsymbol{k=i})$ then $(i\leftrightarrow_{n+1}j)(k)=(i\leftrightarrow_{n+1}j)(i)=j=\big(i\leftrightarrow_n j\big)(i)=(i\leftrightarrow_n j)(k)$

2. $(\boldsymbol{k=j})$ then $(i \leftrightarrow_{n+1} j)(k) = (i \leftrightarrow_{n+1} j)(j) = i = \left(i \leftrightarrow_n j\right)(j) = (i \leftrightarrow_n j)(k)$
3. $(\boldsymbol{k \neq i, j})$ then $(i \leftrightarrow_{n+1} j)(k) = k = (i \leftrightarrow_n j)(k)$

so $(i \leftrightarrow_{n+1} j)_{\{0,\ldots,n\}} = (i \leftrightarrow_n j)$ □

Lemma 10.18. *If $n \in \mathbb{N}_0$ and $i, j \in \{0, \ldots, n\}$ then $(i \leftrightarrow_n j) \circ (i \leftrightarrow_n j) = i_{\{0,\ldots,n\}}$*

Proof. We have the following cases for $k \in \{0, \ldots, n\}$

1. $(\boldsymbol{k=i})$ then $[(i \leftrightarrow_n j) \circ (i \leftrightarrow_n j)](k) = (i \leftrightarrow_n j)((i \leftrightarrow_n j)(k)) = (i \leftrightarrow_n j)(j) = i = k$
2. $(\boldsymbol{k=j})$ then $[(i \leftrightarrow_n j) \circ (i \leftrightarrow_n j)](k) = (i \leftrightarrow_n j)((i \leftrightarrow_n j)(k)) = (i \leftrightarrow_n j)(i) = j = k$
3. $(\boldsymbol{k \neq i, j})$ then $[(i \leftrightarrow_n j) \circ (i \leftrightarrow_n j)](k) = (i \leftrightarrow_n j)((i \leftrightarrow_n j)(k)) = (i \leftrightarrow_n j)(k) = k$

proving the lemma. □

The use of commutativity and associativity of abelian group gives for the

Theorem 10.19. *If $\langle A, +\rangle$ is a **abelian semi-group**, $n \in \mathbb{N}_0$, $\{x_i\}_{i \in \{0,\ldots,n\}}$ a family of elements in A, $\sigma: \{0, \ldots, n\} \to \{0, \ldots, n\}$ a permutation then $\sum_{k=0}^{n} x_i = \sum_{k=0}^{n} x_{\sigma(k)}$*

Proof. Intuitive this is very clear, however a exact proof is rather complex. We prove this by mathematical induction so let $B = \{j \in \mathbb{N}_0 | \text{if } j \in \{0, \ldots, n\} \text{ and } \sigma: \{0, \ldots, j\} \to \{0, \ldots, j\} \text{ a permutation then } \sum_{k=0}^{j} x_{\sigma(k)} = \sum_{k=0}^{j} x_k\}$. We have then

1. Every permutation $\sigma: \{0, \ldots, 0\} \to \{0, \ldots, 0\}$ is equal to $i_{\{0\}}$ and $\sum_{k=0}^{0} x_k = x_0 = x_{\sigma_0} = \sum_{k=0}^{0} x_{\sigma(k)}$ so $0 \in B$
2. If $j \in B$ then for $j+1$ we have the following possibilities to consider
   a. $(\boldsymbol{j+1 \notin \{0, \ldots, n\}})$ then $j+1 \in B$
   b. $(\boldsymbol{j+1 \in \{0, \ldots, n\}})$ then as $0 < j+1 \Rightarrow 1 \leqslant j+1$ we have again the following possibilities to consider
      i. $(\boldsymbol{1 = j+1})$ then we have for the permutation $\sigma: \{0, \ldots, 1\} \to \{0, \ldots, 1\}$ the following possibilities
         A. $(\boldsymbol{\sigma = i_{\{0,1\}}})$ Then $\sum_{k=0}^{1} x_{\sigma(k)} = (\sum_{k=0}^{0} x_{\sigma(k)}) + x_{\sigma(1)} = x_{\sigma(0)} + x_{\sigma(0)} = x_0 + x_1 = (\sum_{k=0}^{0} x_k) + x_1 = \sum_{k=0}^{1} x_k$
         B. $(\boldsymbol{\sigma = (0 \leftrightarrow_1 1)})$ Then $\sum_{k=0}^{1} x_{\sigma(k)} = (\sum_{k=0}^{0} x_{\sigma(k)}) + x_{\sigma(1)} = x_{\sigma(0)} + x_{\sigma(1)} = x_1 + x_0 \underset{\text{commutativity}}{=} x_0 + x_1 = (\sum_{k=0}^{0} x_k) + x_1 = \sum_{k=0}^{1} x_k$

      so in all cases we have $j+1 \in B$

ii. **$(1<j+1)$** Let now $\sigma:\{0,\ldots,j+1\}\to\{0,\ldots,j+1\}$ a permutation then we have for σ either

A. **$(\sigma(j+1)=j+1)$** Using 10.15 we have that $\sigma_{|\{0,\ldots,j\}}:\{0,\ldots,j\}\to\{0,\ldots,j\}$ is a permutation and then $\sum_{k=0}^{j+1}x_{\sigma(k)}=(\sum_{k=0}^{j}x_{\sigma(k)})+x_{\sigma(j+1)}=(\sum_{k=0}^{j}((x_\sigma)_{|\{0,\ldots,j\}})(k))+x_{\sigma(j+)}$, as it is trivially proved that $(x\circ\sigma)_{|\{0,\ldots,j\}}=x\circ\sigma_{|\{0,\ldots,j\}}$ we have $\sum_{k=0}^{j+1}x_k=(\sum_{k=0}^{j}x_{(\sigma_{|\{0,\ldots,j\}})(k)})+x_{\sigma(j+1)}\underset{j\in B}{=}(\sum_{k=0}^{j}x_k)+x_{\sigma(j+1)}\underset{\sigma(j+1)=j+1}{=}(\sum_{k=0}^{j}x_k)+x_{j+1}=\sum_{k=0}^{j+1}x_k$ proving that $j+1\in B$

B. **$(\sigma(j+1)\neq j+1)$** In this case there must by surjectivity of σ exists a $k\in\{1,\ldots j+1\}$ such that $\sigma(k)=j+1$, as $\sigma(j+1)\neq j+1$ we must have then that $k\in\{1,\ldots,j\}$. We have then using the fact that $(j\leftrightarrow_j k):\{0,\ldots,j\}\to\{0,\ldots,j\}$ is a bijection and $j\in B$

$$\begin{aligned}\sum_{i=0}^{j+1}x_{\sigma(i)} &= \sum_{i=0}^{j}x_{\sigma(i)}+x_{\sigma(j+1)}\\ &= \sum_{i=0}^{j}x_{\sigma(j\leftrightarrow_j k)(i)}+x_{\sigma(j+1)}\\ &= \left(\sum_{i=0}^{j-1}x_{\sigma(j\leftrightarrow_j k)(i)}+x_{\sigma(j\leftrightarrow_j k)(j)}\right)+x_{\sigma(j+1)}\\ &= \left(\sum_{i=0}^{j-1}x_{\sigma(j\leftrightarrow_j k)(i)}+x_{\sigma(k)}\right)+x_{\sigma(j+1)}\\ &= \left(\sum_{i=0}^{j-1}x_{\sigma(j\leftrightarrow_j k)(i)}+x_{j+1}\right)+x_{\sigma(j+1)}\\ &= \left(\sum_{i=0}^{j-1}x_{\sigma(j\leftrightarrow_j k)(i)}+x_{\sigma(j+1)}\right)+x_{j+1}\end{aligned}$$

(in last step we used associativity and commutativity). Define now $\gamma:\{0,\ldots,j\}\to\{0,\ldots,j\}$ defined by $i\to\gamma(i)=\begin{cases}\sigma\circ(j\leftrightarrow_j k)(i)\text{ if }i\in\{0,\ldots,j-1\}\\ \sigma_{j+1}\text{ if }i=j\end{cases}$ then we have using the above that

$$\begin{aligned}\sum_{i=0}^{j+1}x_{\sigma(i)} &= \left(\sum_{i=0}^{j-1}x_{\gamma(i)}+x_{\gamma(j)}\right)+x_{j+1}\\ &= \sum_{i=0}^{j}x_{\gamma(i)}+x_{j+1}\end{aligned}$$

If we now prove that γ s bijective and thus a permutation then we have

$$\begin{aligned}\sum_{i=0}^{j+1} x_{\sigma(i)} &= \sum_{i=0}^{j} x_i + x_{j+1}\\ &= \sum_{i=0}^{j+1} x_i\end{aligned}$$

proving that $j+1 \in B$. To prove that γ is bijective consider

1. **(injective)** If $\gamma(i) = \gamma(i')$ then we have either

   a. $\boldsymbol{(i = j = i')}$ then $i = i'$

   b. $\boldsymbol{(i = j, i' \neq j)}$ then $\sigma(j+1) = \gamma(i) = \gamma(i') = \sigma((j \leftrightarrow_j k)(i')) \underset{\sigma\,\text{is injective}}{\Rightarrow} j+1 = (j \leftrightarrow_j k)(i') \in \{0, \ldots, j\}$ giving a contradiction, this case does not apply.

   c. $\boldsymbol{(i \neq j, i' = j)}$ then $\sigma\,(j+1) = \gamma(i') = \gamma(i) = \sigma((j \leftrightarrow_j k)(i)) \underset{\sigma\,\text{is injective}}{\Rightarrow} j+1 = (j \leftrightarrow_j k)(i) \in \{0, \ldots, j\}$ a contradiction and thus this case does not apply.

   d. $\boldsymbol{(i, i' \neq j)}$ then $\sigma((j \leftrightarrow_j k)(i)) = \sigma((j \leftrightarrow_j k)(i')) \underset{\sigma\,\text{is injective}}{\Rightarrow} (j \leftrightarrow_j k)(i) = (j \leftrightarrow_j k)(i') \underset{(i \leftrightarrow_j k)\,\text{is injective}}{\Rightarrow} i = i'$

2. **(surjective)** Let $l \in \{0, \ldots, j\}$ then by surjectivity of σ we have that $\exists m \in \{0, \ldots, j+1\}$ such that $l = \sigma(m)$. We have then for m the following cases

   a. $\boldsymbol{(m = j+1)}$ then $\gamma(j) = \sigma(j+1) = l \Rightarrow j \in \{0, \ldots, j\}$ and $\gamma(j) = l$

   b. $\boldsymbol{(m = j)}$ then we can not have $j = k$ as then $j+1 = \sigma(k) = \sigma(j) = \sigma(m) \in \{0, \ldots, j\}$ a contradiction so we must have $j \neq k$. So $k \in \{0, \ldots, j-1\}$ and $\gamma(k) = \sigma((j \leftrightarrow_j k)(k)) = \sigma(j) = \sigma(m) = l \Rightarrow k \in \{0, \ldots, j\}$ and $\gamma(k) = l$

   c. $\boldsymbol{(m \in \{0, \ldots, j-1\})}$ then as $m \neq k$ [other wise we would have $j+1 = \sigma(m) = l \in \{0, \ldots, j\}$ a contradiction] we have $\gamma(m) = \sigma((j \leftrightarrow_j k)(m)) = \sigma(m) = l$ and thus $m \in \{0, \ldots, j\}$ and $\gamma(m) = l$

□

Definition 10.20. *If* $n, m \in \mathbb{N}_0$ *with* $n \leqslant m$ *then* $\{n, \ldots, m\} = \{x \in \mathbb{N}_0 | n \leqslant x \leqslant m\}$

Definition 10.21. *If* $\langle A, + \rangle$ *is a semi-group,* $n, m \in \mathbb{N}_0$ *with* $n \leqslant m$ *and* $\{x_i\}_{i \in \{n, \ldots, m\}}$ *a family in* A *then we define* $\sum_{i=n}^{m} x_i$ *to be equal to* $\sum_{i=0}^{m-n} x_{i+n}$.

Theorem 10.22. *If* $\langle A, + \rangle$ *is a semi-group* $n, m, k \in \mathbb{N}_0$ *with* $n \leqslant m$ *and* $\{x_i\}_{i \in \{n+k, \ldots, m+k\}}$ *then* $\sum_{i=n}^{m} x_{i+k} = \sum_{i=n+k}^{m+k} x_i$

Proof. $\sum_{i=n}^{m} x_{i+k} \underset{\text{definition}}{=} \sum_{i=0}^{m-n} x_{(i+n)+k} = \sum_{i=0}^{m-n} x_{i+(n+k)} \underset{\text{definition}}{=} \sum_{i=n+k}^{(m-n)+n+k} x_i = \sum_{i=n+k}^{m+k} x_i$ □

Lemma 10.23. *If* $\langle A, + \rangle$ *is a semi-group* $n, m \in \mathbb{N}_0$ *with* $n < m + 1$ *and* $\{x_i\}_{i \in \{n, \ldots, m+1\}}$ *then*

1. $\sum_{i=n}^{n} x_i = x_n$
2. $\sum_{i=n}^{m+1} x_i = \sum_{i=n}^{m} x_i + x_{m+1}$

Proof.

1. $\sum_{i=n}^{n} x_i = \sum_{i=0}^{0} x_{i+n} = x_{0+n} = x_n$
2. $\sum_{i=n}^{m+1} x_i = \sum_{i=0}^{(m-n)+1} x_{i+n} = \sum_{i=0}^{(m-n)} x_{i+n} + x_{(m-n)+1+n} = \sum_{i=n}^{m} x_i + x_{m+1}$ □

Example 10.24. Let $\langle A, + \rangle$ be a semi group, $n, m \in \mathbb{N}_0$ with $n < m$ and $\{x_i\}_{i \in \{n, \ldots, m\}}$ be such that $\forall i \in \{n, \ldots, m\}$ we have $x_i = 0$ then $\sum_{i=n}^{m} x_i = 0$

Proof. We prove this by induction, so let $S = \{i \in \mathbb{N}_0 | \text{If } \{x_i\}_{i \in \{n, \ldots, n+i\}} \text{ is such that } \forall i \in \{n, \ldots, n+i\} \text{ we have } x_i = 0 \text{ then } \sum_{i=n}^{n+i} x_i = 0\}$ then we have

$0 \in \mathcal{S}$. then $\sum_{i=n}^{n+0} x_i = x_n = 0$ proving that $0 \in \mathcal{S}$

$i \in \mathcal{S} \Rightarrow i \in \mathcal{S}$. take then $\sum_{i=n}^{n+(i+1)} x_i = (\sum_{i=n}^{n+i} x_i) + x_{n+(i+1)} \underset{i \in \mathcal{S}}{=} 0 + 0 = 0$ proving that $i + 1 \in \mathcal{S}$

Using mathematical induction we have then that $\boldsymbol{S} = \mathbb{N}_0$ so that $m - n \in \mathcal{S}$ and thus $0 = \sum_{i=n}^{(m-n)+n} x_i = \sum_{i=n}^{m} x_i$. □

Note 10.25. If $\langle A, + \rangle$ is a semi-group $n \in \mathbb{N}_0$ and $k, l \in \{0, \ldots, n\}, k \leqslant l$ then we define $\sum_{i=k}^{l} x_i \equiv \sum_{i=k}^{l} (x_{|\{k, \ldots, l\}})_i$

Theorem 10.26. *Let* $\langle A, + \rangle$ *be a semi-group,* $n \in \mathbb{N} = \mathbb{N}_0 \backslash \{0\}$, $\{x_i\}_{i \in \{0, \ldots, n\}}$ *a family of element of* A *and* $m \in \{0, \ldots, n-1\}$ *then* $\sum_{i=0}^{n} x_i = (\sum_{i=0}^{m} x_i) + (\sum_{i=m+1}^{n} x_i)$

Proof. We prove this by induction on n so let $B=\{n\in\mathbb{N}_0|\text{if }\{x\}_{i\in\{0,\ldots,n+1\}}$ *a* family in A and $m\in\{0,\ldots,n\}$ then $\sum_{i=0}^{n+1} x_i=(\sum_{i=0}^{m} x_i)+(\sum_{i=m+1}^{n+1} x_i)\}$ then we have

1. $\sum_{i=0}^{0+1} x_i=\sum_{i=0}^{1} x_1=(\sum_{i=0}^{0} x_i)+x_1=(\sum_{i=0}^{0} x_i)+x_{0+1}=(\sum_{i=0}^{0} x_i)+(\sum_{i=0}^{0} x_{i+1})=(\sum_{i=0}^{0} x_i)+(\sum_{i=1}^{1} x_i)$ and for $m\in\{0,\ldots,0\}=\{0\}$ we have $m=0\Rightarrow\sum_{i=0}^{0+1} x_i=(\sum_{i=0}^{m} x_i)+(\sum_{i=m+1}^{0+1})$ proving that $0\in B$

2. If $n\in B$ take then $n+1$ and $m\in\{0,\ldots,(n+1)\}$ then we have the following cases for m

   a. $\boldsymbol{(m=n+1)}$ Then $(\sum_{i=0}^{m} x_i)+\left(\sum_{i=m+1}^{(n+1)+1} x_i\right)=(\sum_{i=0}^{n+1} x_i)+\left(\sum_{i=(n+1)+1}^{(n+1)+1} x_i\right)=(\sum_{i=0}^{n+1} x_i)+\left(\sum_{i=0}^{0} x_{i+((n+1)+1)}\right)=(\sum_{i=0}^{n+1} x_i)+x_{(n+1)+1}=\sum_{i=0}^{(n+1)+1} x_i$ proving that $n+1\in B$

   b. $\boldsymbol{(m\in\{0,\ldots,n\})}$ then $(\sum_{i=0}^{m} x_i)+\left(\sum_{i=m+1}^{(n+1)+1} x_i\right)=(\sum_{i=0}^{m} x_i)+\left(\sum_{i=m+1}^{(n+1)+1}(x_{|\{m+1,\ldots,(n+1)+1\}})_i\right)=(\sum_{i=0}^{m} x_i)+\left(\sum_{i=0}^{(((n+1)+1)-(m+1))}(x_{|\{m+1,\ldots,(n+1)+1\}})_{i+(m+1)}\right)=(\sum_{i=0}^{m} x_i)+\left(\sum_{i=0}^{((n+1)-(m+1))+1}(x_{|\{m+1,\ldots,(n+1)+1\}})_{i+(m+1)}\right)=(\sum_{i=0}^{m} x_i)+\left(\left(\sum_{i=0}^{((n+1)-(m+1))}(x_{|\{m+1,\ldots,(n+1)+1\}})_{i+(m+1)}\right)+(x_{|\{m+1\},\ldots,(n+1)+1})_{(((n+1)-(m+1))+1)+(m+1)}\right)=(\sum_{i=0}^{m} x_i)+\left(\left(\sum_{i=m+1}^{(n+1)}(x_{|\{m+1,\ldots,(n+1)+1\}})_i\right)+x_{(n+1)+1}\right)\underset{\text{associativity}}{=}\left((\sum_{i=0}^{m} x_i)+\left(\sum_{i=m+1}^{(n+1)}(x_{|\{m+1,\ldots,(n+1)+1\}})_i\right)\right)+x_{(n+1)+1}=\left((\sum_{i=0}^{m} x_i)+\left(\sum_{i=m+1}^{(n+1)} x_i\right)\right)+x_{(n+1)+1}\underset{n\in B}{=}(\sum_{i=0}^{n+1} x_i)+x_{(n+1)+1}=\sum_{i=0}^{(n+1)+1} x_i$ proving that $n+1\in B$

   So in case (a) and (b) we have $n+1\in B$

Using mathematical induction (see 4.10) we have $B=\mathbb{N}_0$. So if $n\in\mathbb{N}$ then $n=(n-1)+1$ where $n-1\in\mathbb{N}_0=B$ and thus if $m\in\{0,\ldots,n-1\}$ that $\sum_{i=0}^{n} x_i=\sum_{i=0}^{(n-1)+1} x_i\underset{(n-1)\in B}{=}\sum_{i=0}^{m} x_i+\sum_{m+1}^{(n-1)+1} x_i=\sum_{i=0}^{m} x_i+\sum_{i=m+1}^{n} x_i$ □

Theorem 10.27. *Let $\langle A,+\rangle$ be a abelian group then if $n,m\in\mathbb{N}_0$ with $n\leqslant m$ then if $\{x_i\}_{i\in\{n,\ldots,m+1\}}$ is a family in A we have $\sum_{i=n}^{m}(x_{i+1}-x_i)=x_{m+1}-x_n$*

Proof. We prove this by induction so let $\mathcal{S}_n=\left\{k\in\mathbb{N}_0|\text{if }\{x_i\}_{i\in\{n,\ldots,(n+k)+1\}}\text{ is a family in }A\text{ then }\sum_{i=n}^{(n+k)}(x_{i+1}-x_i)=x_{(n+k)+1}-x_i\right\}$ then we have:

1. If $k=0$ then for $\{x_i\}_{i\in\{n,\ldots,(n+0)+1\}}=\{x_i\}_{i\in\{n,\ldots,n+1\}}$ we have $\sum_{i=n}^{n}(x_{i+1}-x_i)=x_{n+1}-x_n$ proving that $0\in S_n$

2. If $k\in\mathcal{S}_n$ then we have if $\{x_i\}_{i\in\{n,\ldots,(n+(k+1))+1\}}$ that $\sum_{i=n}^{(n+(k+1))}(x_{i+1}-x_i)=\sum_{i=n}^{(n+k)+1}(x_{i+1}-x_i)=\sum_{i=n}^{n+k}(x_{i+1}-x_i)+(x_{((n+k)+1)+1}-x_{((n+k)+1)})\underset{k\in\mathcal{S}_n}{=}(x_{(n+k)+1}-x_n)+(x_{(n+(k+1))+1}-x_{(n+k)+1})=(x_{(n+(k+1))+1}-x_n)$ proving that $k+1\in\mathcal{S}_n$

Using mathematical induction we have then that $\mathcal{S}_n = \mathbb{N}_0$ so if $n, m \in \mathbb{N}_0$ with $n \leqslant m$ then $k = m - n \in \mathbb{N}_0 = \mathcal{S}_n$ so that $\sum_{i=n}^{m} (x_{i+1} - x_i) = \sum_{i=n}^{(n+k)} (x_{i+1} - x_i) = x_{(n+k)+1} - x_n = x_{m+1} - x_n$ $\square$

Theorem 10.28. *Let $\langle A, + \rangle$ be a abelian group then if $n, m \in \mathbb{N}_0$ with $n \leqslant m$ then if $\{x_i\}_{i \in \{n, \ldots, m+1\}}$ is a family in A we have $\sum_{i=n}^{m} (x_i - x_{i+1}) = x_n - x_{m+1}$*

Proof. We prove this by induction so let $\mathcal{S}_n = \left\{k \in \mathbb{N}_0 | \text{if } \{x_i\}_{i \in \{n, \ldots, (n+k)+1\}} \text{ is } a \text{ family in } A \text{ then } \sum_{i=n}^{(n+k)} (x_i - x_{i+1}) = x_n - x_{(n+k)+1}\right\}$ then we have:

1. If $k = 0$ then for $\{x_i\}_{i \in \{n, \ldots, (n+0)+1\}} = \{x_i\}_{i \in \{n, \ldots, n+1\}}$ we have $\sum_{i=n}^{n} (x_i - x_{i+1}) = x_n - x_{n+1}$ proving that $0 \in S_n$
2. If $k \in \mathcal{S}_n$ then we have if $\{x_i\}_{i \in \{n, \ldots, (n+(k+1))+1\}}$ that $\sum_{i=n}^{(n+(k+1))} (x_i - x_{i+1}) = \sum_{i=n}^{(n+k)+1} (x_i - x_{i+1}) = \sum_{i=n}^{n+k} (x_i - x_{i+1}) + (x_{((n+k)+1)} - x_{((n+k)+1)+1}) \underset{k \in \mathcal{S}_n}{=} (x_n - x_{(n+k)+1}) + (x_{n+(k+1)} - x_{((n+k)+1)+1}) = (x_n - x_{(n+(k+1))+1})$ proving that $k + 1 \in \mathcal{S}_n$

Using mathematical induction we have then that $\mathcal{S}_n = \mathbb{N}_0$ so if $n, m \in \mathbb{N}_0$ with $n \leqslant m$ then $k = m - n \in \mathbb{N}_0 = \mathcal{S}_n$ so that $\sum_{i=n}^{m} (x_i - x_{i+1}) = \sum_{i=n}^{(n+k)} (x_i - x_{i+1}) = x_n - x_{(n+k)+1} = x_n - x_{m+1}$ $\square$

Theorem 10.29. (Generalized Associativity) *Let $\langle A, + \rangle$ be a semi-group. Take $\{(b_j, e_j)\}_{j \in \{0, \ldots, m\}}$ a sequence of natural number pairs such that $b_0 = 0$, $e_m = n$, $\forall j \in \{0, \ldots, m\}$ we have $b_j \leqslant e_j$ and $\forall j \in \{0, \ldots, m-1\}$ we have $e_j + 1 = b_{j+1}$. Then we have for a family $\{x_i\}_{i \in \{0, \ldots, n\}}$ of element in A that $\sum_{i=0}^{n} x_i = \sum_{k=0}^{m} (\sum_{l=b_k}^{e_k} x_l)$*

Proof. We prove this by induction on m so take $B = \{m \in \mathbb{N}_0 | \text{let } \{(b_j, e_j)\}_{j \in \{0, \ldots, m\}}$ be a sequence of natural number pairs such that $b_0 = 0$, $e_m = n$, $\forall j \in \{0, \ldots, m\}$ we have $b_j \leqslant e_j$ and $\forall j \in \{0, \ldots, m\}$ we have $e_j + 1 = b_{j+1}$ .If $\{x_i\}_{i \in \{0, \ldots, n\}}$ is a family of elements in A then $\sum_{i=0}^{n} x_i = \sum_{k=0}^{m} (\sum_{l=b_k}^{e_k} x_l)\}$. We have then

1. If $m = 0$ then $b_0 = 0, e_0 = e_m = n$ and $\sum_{i=0}^{n} x_i = \sum_{l=b_0}^{e_0} x_l = \sum_{k=0}^{m} (\sum_{l=b_k}^{e_k} x_l)$ so $0 \in B$
2. If $m \in B$ take then $m+1$ and consider $\{(b_j, e_j)\}_{j \in \{0, \ldots, m+1\}}$ with $b_0 = 0$, $e_{m+1} = n$, $\forall j \in \{0, \ldots, m+1\} \vDash b_j \leqslant e_j$ and $\forall j \in \{0, \ldots, m\} \vDash e_j + 1 = b_{j+1}$, consider also $\{x_i\}_{i \in \{0, \ldots, n\}}$ of elements in A. We have then if $e_m = p$ that for $\{(b_j, e_j)\}_{j \in \{0, \ldots, m\}}$ we still have $b_0 = 0, e_m = p$ and $\forall j \in \{0, \ldots, m\} \vDash b_j \leqslant e_j$ and $\forall j \in \{0, \ldots, m-1\} \vDash e_j + 1 = b_{j+1}$. Using the fact that $m \in B$ we have then that $\sum_{i=0}^{p} x_i = \sum_{k=0}^{m} (\sum_{l=b_k}^{e_k} x_l)$. Now as $p < p+1 = e_m + 1 = b_{m+1} \leqslant e_{m+1} = n$ we have $p \in \{0, \ldots, n-1\}$ and thus by 10.26 we have $\sum_{i=0}^{n} x_i = (\sum_{i=0}^{p} x_i) + (\sum_{i=p+1}^{n} x_i) = (\sum_{k=0}^{m} (\sum_{l=b_k}^{e_k} x_l)) + (\sum_{l=b_{m+1}}^{e_{m+1}} x_l) = \sum_{k=0}^{m+1} (\sum_{l=b_k}^{e_k} x_l) \Rightarrow m+1 \in B$

So by mathematical induction we have $B = \mathbb{N}_0$ proving our theorem. $\square$

Theorem 10.30. *Let $\langle F, +, \cdot\rangle$ be a field, $\alpha \in F$ and $\{x_i\}_{i\in\{0,\ldots,n\}}$ a family in F then $\sum_{i=0}^{n} (\alpha \cdot x_i) = \alpha \cdot \sum_{i=0}^{n} x_i$*

Proof. This is proved by induction so let $S = \{n \in \mathbb{N}_0 | \text{if } \{x_i\}_{i\in 0,\ldots n} \text{ is } a \text{ family in } F \text{ and } \alpha \in F \text{ then } \sum_{i=0}^{n} \alpha \cdot x_i = \alpha \cdot \sum_{i=0}^{n} x_i\}$ then we have

1. If $n = 0$ then $\sum_{i=0}^{0} a \cdot x_i = \alpha \cdot x_o = \alpha \cdot \sum_{i=1}^{0} x_i$ proving that $0 \in S$
2. If $n \in S$ then we have $\sum_{i=0}^{n+1} \alpha \cdot x_i = \sum_{i=0}^{n} \alpha \cdot x_i + \alpha \cdot x_{n+1} \underset{n\in S}{=} \alpha \cdot \sum_{i=0}^{n} x_i + \alpha \cdot x_{n+1} = \alpha \cdot (\sum_{i=0}^{n} x_i + x_{n+1}) = \alpha \cdot \sum_{i=0}^{n+1} x_i$ proving that $n + 1 \in S$

Mathematical induction completes then the theorem. □

Theorem 10.31. *Let $\mathbb{R}$ be the real field, $n \in \mathbb{N}_0$ and $\{c\}_{i\in\{0,\ldots,n\}}$ be constant family in $\mathbb{R}$ then we have that $\sum_{i=0}^{n} c = (n+1) \cdot c$*

Proof. Then we have

1. If $n = 0$ then $\sum_{i=0}^{n} c = c = 1 \cdot n = (0+1) \cdot n$ proving that $0 \in S$
2. If $n \in S$ then $\sum_{i=0}^{n+1} c = \sum_{i=0}^{n} c + c \underset{n\in S}{=} (n+1) \cdot c + c = ((n+1)+1) \cdot c$ proving that $n + 1 \in S$ □

Theorem 10.32. *Let $\mathbb{R}$ (or $\mathbb{R}_{\mathbb{C}}$) be the real field (embedded real field in the complex numbers, $\{x_i\}_{i\in\{0,\ldots,n\}}$ a family in $\mathbb{R}$ (or $\mathbb{R}_{\mathbb{C}}$) such that $\forall i \in \{0, \ldots, n\}$ we have $0 \leqslant x_i$ then $0 \leqslant (<) \sum_{i=0}^{n} x_i$ (where we have $<$ if $\exists i \in \{0, \ldots, n\}$ such that $0 < x_i$)*

Proof. We prove this by induction so let $S = \{n \in \mathbb{N} | \text{if } \{x_i\}_{i\in\{0,\ldots,n\}} \text{ is } a \text{ family of non negative numbers then } 0 \leqslant (<) \sum_{i=0}^{n} x_i \text{ where we have} < \text{if at least one of the } x_i's \text{ is} > 0\}$, we have then

1. if $n = 0$ then $0 \leqslant x_0 = \sum_{i=0}^{0} x_i$ and if $\exists i \in (0, \ldots, 0)$ such that $0 < x_i = x_0$ then we have $0 < x_0 = \sum_{i=0}^{0} x_i$ proving that $0 \in S$
2. If $n \in S$ then we have $\sum_{i=0}^{n+1} x_i = \sum_{i=0}^{n} x_i + x_{n+1} \geqslant \sum_{i=0}^{n} x_i + 0 \geqslant_{n\in S} 0$. If now $\exists i \in \{0, \ldots, n+1\}$ such that $x_i > 0$ then we have two cases to consider
   a. $\boldsymbol{(i \in \{0, \ldots, n\})}$ then as $n \in S$ we have $0 < \sum_{i=0}^{n} x_i$ so that $\sum_{i=0}^{n+1} x_i \geqslant \sum_{i=0}^{n} x_i > 0$
   b. $\boldsymbol{(i = n + 1)}$ then we have $\sum_{i=0}^{n+1} x_i = \sum_{i=0}^{n} x_i + x_{n+1} > \sum_{i=0}^{n} x_i + 0 \geqslant_{n\in S} 0$

   This proves that $n + 1 \in S$

By mathematical induction we have then that $S = \mathbb{N}_0$ proving the theorem □

Corollary 10.33. *Let $\mathbb{R}$ (or $\mathbb{R}_{\mathbb{C}}$) be the real field (embedded real field in the complex numbers, $\{x_i\}_{i\in\{0,\ldots,n\}}, \{y_i\}_{i\in\{0,\ldots,n\}}$ a family in $\mathbb{R}$ (or $\mathbb{R}_{\mathbb{C}}$) such that $\forall i \in \{0, \ldots, n\}$ we have $x_i \leqslant y_i$ then $\sum_{i=0}^{n} x_i \leqslant (<) \sum_{i=0}^{n} y_i$ (where we have $<$ if $\exists i \in \{0, \ldots, n\}$ such that $x_i < y_i$)*

Proof. Take the family $\{y_i+(-x_i)\}_{i\in\{0,\ldots,n\}}$ then $\forall i\in\{0,\ldots,n\}$ we have $x_i+(-y_i)\geqslant 0$ so by 10.32 we have

$$\sum_{i=0}^{n}\,(y_i+(-x_i))\geqslant(>)0\,(>\text{if we have one } y_i>x_i) \tag{10.1}$$

Now $\sum_{i=0}^{n}\,(y_i+(-x_i))\underset{10.11}{=}\sum_{i=0}^{n}\,y_i+\sum_{i=0}^{n}\,(-x_i)\underset{10.30}{=}\sum_{i=0}^{n}\,y_i-\sum_{i=0}^{n}\,x_i$ which together with 10.1 we have $\sum_{i=0}^{n}y_i\geqslant(>)\sum_{i=0}^{n}x_i$ ($>$ if we have one $y_i>x_i$) $\square$

Definition 10.34. (Support) *If $\langle A,+\rangle$ is a semi-group and $\{x_i\}_{i\in I}$ a family of elements in A. Then* $\operatorname{support}(\{x_i\}_{i\in I})=\{i\in I|x_i\neq 0\}$ *where* 0 *is the neutral element*

Definition 10.35. (Finite support) *If $\langle A,+\rangle$ is a semi-group and $\{x_i\}_{i\in I}$ a family of elements in A has finite support if* $\operatorname{support}(\{x_i\}_{i\in I})$ *is finite.*

Example 10.36. If $\langle A,+\rangle$ is a semi-group and $\{x_i\}_{i\in I}$ is a family of elements in A where I is finite then as $\operatorname{support}(\{x_i\}_{i\in I})\subseteq I$ we have by 5.45 that $\operatorname{support}(\{x_i\}_{i\in I})$ is finite and thus $\{x_i\}_{i\in I}$ has finite support.

Definition 10.37. (generalized sum) *Let $\langle A,+\rangle$ be a abelian semi-group and $\{x_i\}_{i\in I}$ a family of elements in A with finite support. Then $\sum_{i\in I}x_i$ is defined to be*

1. *If* $\operatorname{support}(\{x_i\}_{i\in I})=\emptyset$ *then* $\sum_{i\in I}x_i=0$ *(neutral element of the semi-group)*
2. *If* $\operatorname{support}(\{x_i\}_{i\in I})\neq\emptyset$ *then* $\sum_{i\in I}x_i=\sum_{i=0}^{n-1}x_{\beta(i)}$ *where* $\beta\colon\{0,\ldots,n-1\}=S_n\rightarrow\operatorname{support}(\{x_i\}_{i\in I})$ *is a bijection (which must exists because of finiteness of the support)*

We must of course prove that (2) of the definition is independent of the choosen bijection

Proof. As $\operatorname{support}(\{x_i\}_{i\in I})$ is finite there exists a unique (see 5.42) $n\in\mathbb{N}_0$ such that $n=S_n\approx\operatorname{support}(\{x_i\}_{i\in I})$, which as $\operatorname{support}(\{x_i\}_{i\in I})\neq\emptyset$ means that $n>0$. If $\beta'\colon\{0,\ldots,n-1\}\rightarrow I$ is another bijection then we have that $\tau=\beta'^{-1}\circ\beta\colon\{0,\ldots,n-1\}\rightarrow\{0,\ldots,n-1\}$ is also a bijection and $\sum_{i=0}^{n-1}x_{\beta'(i)}\underset{10.19}{=}\sum_{i=0}^{n-1}x_{\beta'((\beta'^{-1}\circ\beta)(i))}=\sum_{i=0}^{n-1}x_{\beta'(\beta'^{-1}(\beta(i)))}=\sum_{i=0}^{n-1}x_{\beta(i)}$ $\square$

Example 10.38. Let $\langle A,+\rangle$ be a abelian semi-group and $\{x_i\}_{i\in I}\subseteq A$ such that $\forall i\in I$ we have $x_i=0$ then $\sum_{\in I}x_i=0$

Proof. By the hypothese we have that $\sup(\{x_i\}_{i\in I})=\emptyset$ so that by definition $\sum_{i\in I}x_i=0$ $\square$

We have for a multiplicative group a equivalent definition, only we note then the neutral element with 1 and use $\Pi_{i\in I}x_i$ instead of the summation symbol all the following theorems are valid for a multiplicative group if we do this substitution. So we use then the following definition:

Definition 10.39. (generalized product) *Let $\langle A, +\rangle$ be a multiplicative abelian semi-group and $\{x_i\}_{i\in I}$ a family of elements in A with finite support [*$\text{support}(A)=\{x\in A|x\neq 1\}$*] Then $\prod_{i\in I} x_i$ is defined to be*

1. *If* $\text{support}(\{x_i\}_{i\in I})=\emptyset$ *then $\prod_{i\in I} x_i=1$ (neutral element of the semi-group)*
2. *If* $\text{support}(\{x_i\}_{i\in I})\neq\emptyset$ *then $\prod_{i\in I} x_i=\prod_{i=0}^{n-1} x_{b(i)}$ where $\beta\colon\{0,\ldots,n-1\}=S_n\rightarrow$* $\text{support}(\{x_i\}_{i\in I})$ *is a bijection (which must exists because of finiteness of the support)*

Lemma 10.40. *Let $\langle X, +\rangle$ be a abelian semi-group, I a finite set with $n=\#(I)$ and $\{x_i\}_{i\in I}\subseteq X$ a finite family of elements in X then*

1. *If $I=\emptyset$ then $\sum_{i\in I} x_i=0$*
2. *If $I\neq\emptyset$ then $\sum_{i\in I} x_i=\sum_{i=0}^{n} x_{\beta(i)}$ where $n=\#(I)-1$ and $\beta\colon\{0,\ldots,n\}\rightarrow I$ is a bijection.*
3. *For every bijection $\alpha\colon\{0,\ldots,\#(I)-1\}\rightarrow I$ we have that $\sum_{i\in I} x_i=\sum_{i=0}^{n} x_{\alpha(i)}$*

Proof.

1. If $I=\emptyset$ then $\sup(\{x_i\}_{i\in I})=\emptyset$ so that $\sum_{i\in I} x_i=0$ by definition.
2. Let $S=\text{support}(\{x_i\}_{i\in I})\subseteq I$ then S is finite (so that $\sum_{i\in I} x_i$ is defined) and also $T=I\setminus S=\{i\in I|x_i=0|\}$ is finite. We have then the following cases for S

   S $=\emptyset$**.** then $T=I\setminus S=I\neq 0$ and thus there exists a bijection $\beta\colon\{0,\ldots,n\}\rightarrow I$ where $n=\#(I)-1$ then $\sum_{i=0}^{n} x_i \underset{10.24}{=} 0 \underset{s=\emptyset \text{ and definition}}{=} \sum_{i\in I} x_i$

   S $=$ ***I*****.** then as $I\neq\emptyset$ there exists a bijection $\beta\colon\{0,\ldots,n\}\rightarrow I=S$ where $n=\#(I)-1=\#(S)-1$ and $\sum_{i\in I} x_i \underset{\text{definition}}{=} \sum_{i=0}^{n} x_{\beta(i)}$

   $\emptyset\neq$ ***S*** $\subset$ ***I*****.** then $S, T\neq\emptyset$ then there exissts bijections $\beta_1\colon\{0,\ldots,n_1\}\rightarrow S$ and $\beta_2\colon\{0,\ldots,n_2\}\rightarrow T$ where $n_1=\#(S)-1$, $n_2=\#(T)-1$. Define then $\beta\colon\{0,\ldots,n_1+n_2+1\}\rightarrow I$ by $\beta(i)=\begin{cases}\beta_1(i) \text{ if } i\in\{0,\ldots,n_1\}\\ \beta_2(i-n_1-1) \text{ if } i\in\{n_1+1,\ldots.,n_1+n_2+1\}\end{cases}$ then we prove that β is a bijection

   injectivity. Let $i,j\in\{0,\ldots,n_1+n_2\}$ with $\beta(i)=\beta(j)$. For i we must then consider the following cases:

   $\boldsymbol{i\in\{0,\ldots,n_1\}}$**.** then we can not have that $j\in\{n_1+1,\ldots,n_1+n_2\}$ for then $\beta(i)=\beta_1(i)\in S$ and $\beta(j)=\beta_2(j-n_1-1)\in T$ which as $\beta(i)=\beta(j)$ would mean that $\emptyset\neq S\bigcap T=\emptyset$ a contradiction. So $j\in\{0,\ldots,n_1\}$ and then $\beta_1(i)=\beta(i)=\beta(j)=\beta_1(j) \underset{\beta_1 \text{ is a bijection}}{\Rightarrow} i=j$

   $\boldsymbol{i\in\{n_1+1,\ldots,n_1+n_2+1\}}$**.** then we can not have that $j\in\{0,\ldots,n_1\}$ for then $\beta(i)=\beta_2(i-n_1-1)\in S$ and $\beta(j)=\beta_1(j)\in T$ which as $\beta(i)=\beta(j)$ would mean that $\emptyset\neq S\bigcap T=\emptyset$ a contradiction. So $j\in\{n_1+1,\ldots,n_1+n_2\}$ and then $\beta_2(i)=\beta(i)=\beta(j)=\beta_2(j) \underset{\beta_2 \text{ is a bijection}}{\Rightarrow} i=j$

surjectivity. If $x \in I$ then either $x \in S$ and then $\exists i \in \{0, \ldots, n_1\} \subseteq \{0, \ldots, n_1 + n_2\}$ such that $x = \beta_1(i) = \beta(i)$ or $x \in T$ and then there exists a $i \in \{0, \ldots, n_2\}$ such that $\beta_2(i) = x$ hence for $i + n_1 + 1 \in \{n_1, \ldots, n_1 + n_2 + 1\}$ we have $\beta(i + n_1 + 1) = \beta_2(i) = x$

Finally $\sum_{i=0}^{n_1+n_2+1} x_{\beta(i)} \underset{10.26}{=} \sum_{i=0}^{n_1} x_{\beta(i)} + \sum_{i=n_1+1}^{n_1+n_2+1} x_{\beta(i)} = \sum_{i=0}^{n_1} x_{\beta_1(i)} + \sum_{i=0}^{n_2} x_{\beta(i+n_1+1)} = \sum_{i=0}^{n_1} x_{\beta_1(i)} + \sum_{i=0}^{n_2} x_{\beta_2(i)} = \sum_{i \in I} x_i + 0$ hence if we take $n = n_1 + n_2 + 1$ we have a bijection $\beta: \{0, \ldots, n\} \rightarrow I$ such that $\sum_{i=0}^{n} x_{\beta(i)} = \sum_{i \in I} x_i$

3. If $\alpha: \{0, \ldots, n\} \rightarrow I$ (where $n = \#(I) - 1$) is another bijection then $\beta^{-1} \circ \alpha: \{0, \ldots, \#(I) - 1\}$ is a bijection and we have $\sum_{i \in I} x_i = \sum_{i=0}^{n} x_{\beta(i)} \underset{10.19}{=} \sum_{i=0}^{n} x_{\beta((\beta^{-1} \circ \alpha)(i))} = \sum_{i=0}^{n} x_{\alpha(i)}$ □

Lemma 10.41. *Let $\langle A, + \rangle$ be a abelian semi-group and $\{x_i\}_{i \in \{n, \ldots, m\}}$ a family in A where $n \leqslant m$, $n, m \in \mathbb{N}_0$ then $\sum_{i \in \{n, \ldots, m\}} x_i = \sum_{i=n}^{m} x_i$*

Proof. Let $\beta: \{0, \ldots, m - n\} \rightarrow \{n, \ldots, m\}$ be the bijection defined by $\beta(i) = i + n$ then using 10.40 we have that $\sum_{i \in \{n, \ldots, m\}} x_i = \sum_{i=0}^{m-n} x_{\beta(i)} = \sum_{i=0}^{m-n} x_{i+n} \underset{10.21}{=} \sum_{i=n}^{m} x_i$ □

Theorem 10.42. *Let $\langle A, + \rangle$ be a abelian semi-group and $\{x_i\}_{i \in I}$ a family in A with finite support then if* $\mathrm{support}(\{x_i\}_{i \in I}) \subseteq J \subseteq I$ *we have* $\sum_{i \in I} x_i = \sum_{i \in J} x_i$

Proof. If $k \in \mathrm{support}(\{x_i\}_{i \in I}) \Rightarrow x_k \neq 0 \underset{\mathrm{support}(\{x_i\}_{i \in I} \subseteq J)}{\Rightarrow} x_k \neq 0 \wedge k \in J \Rightarrow k \in \mathrm{support}(\{x_i\}_{i \in J})$, also if $k \in \mathrm{support}(\{x_i\}_{i \in J}) \Rightarrow k \in J \wedge x_k \neq 0 \underset{J \subseteq I}{\Rightarrow} k \in \mathrm{support}(\{x_i\}_{i \in I})$, this proves that $\mathrm{support}(\{x_i\}_{i \in I}) = \mathrm{support}(\{x_i\}_{i \in J})$. So if $b: \{0, \ldots, n - 1\} = S_n \rightarrow \mathrm{support}(\{x_i\}_{i \in I}) = \mathrm{support}(\{x_i\}_{i \in J})$ is a bijection then $\sum_{i \in I} x_i = \sum_{i=0}^{n-1} x_{b_i} = \sum_{i \in J} x_i$. □

Note 10.43. If $\langle A, \cdot \rangle$ is a abelian semi-group, where the operator is interpreted as a product, and $\{x_i\}_{i \in I}$ has finite support then $\sum_{i \in I} x_i$ is written as $\prod_{i \in I} x_i$. In general we use for a product 1 as the unit so if $\mathrm{support}(\{x_i\}_{i \in I}) = \emptyset$ then $\prod_{i \in I} x_i = 1$

Lemma 10.44. *Let $\langle A, + \rangle$ be a abelian semi-group, $\{x_i\}_{i \in I}$ a family of elements in A with finite support and $f: J \rightarrow I$ a bijection then $\{x_{f(i)}\}_{i \in J}$ has finite support given by* $\mathrm{support}(\{x_{f(i)}\}_{i \in J}) = f^{-1}(\mathrm{support}(\{x_i\}_{i \in I}))$ *and* $\sum_{i \in J} x_{f(i)} = \sum_{i \in I} x_i = \sum_{i \in f(J)} x_i$

Proof. First we prove the assertion about the support

$$
\begin{aligned}
j \in \mathrm{support}(\{x_{f_i}\}_{i \in J}) &\Leftrightarrow x_{f_j} \neq 0 \\
&\Leftrightarrow x(f(j)) \neq 0 \\
&\Leftrightarrow f(j) \in \mathrm{support}(\{x_i\}_{i \in I}) \\
&\Leftrightarrow j \in f^{-1}(\mathrm{support}(\{x_i\}_{i \in I}))
\end{aligned}
$$

As $f: J \rightarrow I$ is a bijection we have that $f_{|f^{-1}(\text{support}(\{x_i\}_{i\in I}))}: f^{-1}(\text{support}(\{x_i\}_{i\in I})) = \text{support}(\{x_{f_i}\}_{i\in J}) \rightarrow \text{support}(\{x_i\}_{i\in I})$ is a bijection and thus $\text{support}(\{x_{f_i}\}_{i\in J}) \approx \text{support}(\{x_i\}_{i\in I})$ so support $(\{x_{f_i}\}_{i\in J})$ is also finite and there exists a $n \in \mathbb{N}_0$ such that $S_n \approx \text{support}(\{x_{f_i}\}_{i\in J})$ and $S_n \approx \text{support}(\{x_i\}_{i\in I})$. So if $h: S_n \rightarrow \text{support}(\{x_{f_i}\}_{i\in J})$ is a bijection then $\sum_{j\in J} x_{f_j} = \sum_{j=0}^{n-1} x_{f_{h_i}}$. But also $g = f_{|f^{-1}(\text{support}(\{x_i\}_{i\in I}))} \circ h: S_n \rightarrow \text{support}(\{x_i\}_{i\in I})$ is a bijection so that $\sum_{i\in I} x_i = \sum_{i=0}^{n-1} x_{g_i} = \sum_{i=0}^{n-1} g_{f_{h_i}} = \sum_{j\in J} x_{f_j}$ □

Lemma 10.45. *If $\langle A, +\rangle$ is a abelian semi-group $\{x_i\}_{i\in I}$ a family of elements with finite support and J, K are sets such that $J \bigcap K = \emptyset$ and $I = J \bigcup K$ then $\{x_i\}_{i\in J}$ and $\{x_i\}_{i\in K}$ have finite support and $\sum_{i\in I} x_i = (\sum_{i\in J} x_i) + (\sum_{i\in K} x_i)$*

Proof. First as $J \bigcup K = I \Rightarrow J, K \subseteq I$ now if $i \in \text{support}(\{x_i\}_{i\in J}) \Rightarrow i \in J \wedge x_i \neq 0 \Rightarrow i \in I \wedge x_i \neq 0 \Rightarrow i \in \text{support}(\{x_i\}_{i\in I}) \Rightarrow \text{support}(\{x_i\}_{i\in J}) \subseteq \text{support}(\{x_i\}_{i\in I})$, similar we have $\text{support}(\{x_i\}_{i\in K}) \subseteq \text{support}(\{x_i\}_{i\in I})$. Using 5.45 we have then that $\text{support}(\{x_i\}_{i\in J})$ and $\text{support}(\{x_i\}_{i\in K})$ are finite and thus that $\{x_i\}_{i\in J}$ and $\{x_i\}_{i\in k}$ have finite support. Also we have $\text{support}(\{x_i\}_{i\in J}) \bigcup \text{support}(\{x_i\}_{i\in K}) \subseteq \text{support}(\{x_i\}_{i\in I})$. If $i \in \text{support}(\{x_i\}_{i\in I}) \Rightarrow i \in I \wedge x_i \neq 0 \underset{I=J\bigcup K}{\Rightarrow} (i \in J \vee i \in K) \wedge x_i \neq 0 \Rightarrow (i \in J \wedge x_i \neq 0) \vee (i \in K \wedge x_i \neq 0) \Rightarrow i \in \text{support}(\{x_i\}_{i\in J}) \bigcup \text{support}(\{x_i\}_{i\in K}) \Rightarrow \text{support}(\{x_i\}_{i\in I}) \subseteq \text{support}(\{x_i\}_{i\in J}) \bigcup \text{support}(\{x_i\}_{i\in K})$ and thus we have $\text{support}(\{x_i\}_{i\in J}) \bigcup \text{support}(\{x_i\}_{i\in K}) = \text{support}(\{x_i\}_{i\in I})$. Also if $i \in \text{support}(\{x_i\}_{i\in J}) \bigcap \text{support}(\{x_i\}_{i\in K}) \Rightarrow (i \in J \wedge x_i \neq 0) \wedge (i \in K \wedge x_i \neq 0) \Rightarrow i \in I \wedge i \in J \Rightarrow i \in I \bigcap J = \emptyset$ which is a contradiction, so we must have $\text{support}(\{x_i\}_{i\in J}) \bigcap \text{support}(\{x_i\}_{i\in K}) = \emptyset$ We have then the following cases to consider:

1. **($\text{support}(\{x_i\}_{i\in J}), \text{support}(\{x_i\}_{i\in K}) = \emptyset$)** in this case we have $\text{support}(\{x_i\}_{i\in I}) = \emptyset$ and thus $\sum_{i\in I} x_i = 0 = 0 + 0 = (\sum_{i\in J} x_i) + (\sum_{i\in K} x_i)$

2. **($\text{support}(\{x_i\}_{i\in J}) = \emptyset, \text{support}(\{x_i\}_{i\in K}) \neq \emptyset$)** we must then have $\text{support}(\{x_i\}_{i\in K}) = \text{support}(\{x_{i\in I}\})$ and if $b: \{0, ..., n-1\} \rightarrow \text{support}(\{x_i\}_{i\in I}) = \text{support}(\{x_i\}_{i\in K})$ is a bijection then we have $\sum_{i\in I} x_i = \sum_{i=0}^{n-1} x_{b_i} = \sum_{i\in K} x_i = 0 + (\sum_{i\in K} x_i) = (\sum_{i\in J} x_i) + (\sum_{i\in K} x_i)$

3. **($\text{support}(\{x_i\}_{i\in J}) \neq \emptyset, \text{support}(\{x_i\}_{i\in K}) = \emptyset$)** we must then have $\text{support}(\{x_i\}_{i\in J}) = \text{support}(\{x_{i\in I}\})$ and if $b: \{0, ..., n-1\} \rightarrow \text{support}(\{x_i\}_{i\in I}) = \text{support}(\{x_i\}_{i\in J})$ is a bijection then we have $\sum_{i\in I} x_i = \sum_{i=0}^{n-1} x_{b_i} = \sum_{i\in J} x_i = (\sum_{i\in J} x_i) + 0 = (\sum_{i\in J} x_i) + (\sum_{i\in K} x_i)$

4. **($\text{support}(\{x_i\}_{i\in J}), \text{support}(\{x_i\}_{i\in K}) \neq \emptyset$)** then as the supports are finite there exists bijections $f': \{0, ..., n_1-1\} \rightarrow \text{support}(\{x_i\}_{i\in J})$, $g: \{0, ..., n_2-1\} \rightarrow \text{support}(\{x_i\}_{i\in K})$. Define now $f: \{n_2, ..., (n_1+n_2) - 1\} \rightarrow \text{support}(\{x_i\}_{i\in J})$ by $i \rightarrow f(i) = f'(i-n_2)$ [which is well defined as if $i \in \{n_2, ...(n_1+n_2)-1\} \Rightarrow n_2 \leqslant i \leqslant n_2 + (n_1-1) \Rightarrow 0 \leqslant i - n_2 \leqslant n_1 - 1$] then f is a bijection.

   a. **(injectivity)** if $f(i) = f(j) \Rightarrow f'(i-n_2) = f'(j-n_2) \underset{f' \text{is bijective}}{\Rightarrow} i - n_2 = j - n_2 \Rightarrow i = j$

b. **(surjectivity)** if $j \in \text{support}(\{x_i\}_{i \in J})$ then as f' is a bijection there exists a $i \in \{0, \ldots, n_1 - 1\}$ such that $f'(i) = j$. From $0 \leqslant i \leqslant n_1 - 1$ we have $n_2 \leqslant i + n_2 \leqslant (n_1 + n_2) - 1$ and $f(i + n_2) = f'((i + n_2) - n_2) = f'(i) = j$

Using $\{0, \ldots, n_2 - 1\} \bigcap \{n_2, \ldots, (n_1 + n_2) - 1\} = \emptyset$, $\text{support}(\{x_i\}_{i \in J}) \bigcap \text{support}(\{x_i\}_{i \in K}) = \emptyset$ and 2.43 we have that $g \bigcup f: \{0, \ldots, n_2 - 1\} \bigcup \{n_2, \ldots, (n_1 + n_2) - 1\} = \{0, \ldots, (n_1 + n_2) - 1\} \rightarrow \text{support}(\{x_i\}_{i \in J}) \bigcup \text{support}(\{x_i\}_{i \in K}) = \text{support}(\{x_i\}_{i \in I})$ is a bijection. Now $\sum_{i \in I} x_i = \sum_{i=0}^{(n_1 + n_2) - 1} x_{(g \bigcup f) i} \underset{10.26}{=} (\sum_{i=0}^{n_2 - 1} x_{(g \bigcup f)_i}) + \left(\sum_{i=n_2}^{(n_1 + n_2) - 1} x_{(g \bigcup f)_i}\right) = (\sum_{i=0}^{n_2 - 1} x_{(g \bigcup f)(i)}) + \left(\sum_{i=n_2}^{(n_1 + n_2) - 1} x_{(g \bigcup f)(i)}\right) = (\sum_{i=0}^{n_2 - 1} x_{g(i)}) + \left(\sum_{i=n_2}^{(n_1 + n_2) - 1} x_{f(i)}\right) = (\sum_{i=0} x_{g_i}) + \left(\sum_{i=n_2}^{(n_1 + n_2) - 1} x_{f'(i - n_2)}\right) = (\sum_{i \in J} x_j) + (\sum_{i=0}^{n_1 - 1} x_{f'((i - n_2) + n_2)}) = (\sum_{i \in J} x_j) + (\sum_{i=0}^{n_1 - 1} x_{f'_i}) = (\sum_{i \in J} x_i) + (\sum_{i \in K} x_i)$ □

Theorem 10.46. *Let $\langle A, + \rangle$ be a abelian semi-group, $\{x_i\}_{i \in I}$ a family of elements in A with finite support and $\{I_i\}_{i \in \{0, \ldots, n\}}, n \in \mathbb{N}$ a family of sets in $\mathcal{P}(I)$ such that $\bigcup_{i \in \{0, \ldots, n\}} I_i = I$ and $\forall k, l \in \{0, \ldots, n\}$ we have $I_k \bigcap I_l = \emptyset$ then we have that $\sum_{i \in I} x_i = \sum_{i=0}^{n} \left(\sum_{j \in I_i} x_j\right)$*

Proof. We prove this by mathematical induction on n (see 4.77) so let $X = \{n \in \{1, \ldots\} = \mathbb{N} | \text{If } \{x_i\}_{i \in I}$ is a family of elements in A with finite support and $\{I_i\}_{i \in \{1, \ldots, n\}}$ is a family of subsets of I such that $\bigcup_{i \in \{0, \ldots, n\}} I_i = I$ and $\forall k, l \in \{0, \ldots, n\}$ we have $I_k \bigcap I_l = \emptyset \Rightarrow \sum_{i \in I} x_i = \sum_{i=0}^{n} (\sum_{j \in I_i} x_j)\}$ then we have:

1. If $n = 1$ then $I = I_0 \bigcup I_1$ and $I_0 \bigcap I_1 = \emptyset$ and using the previous lemma we have $\sum_{i \in I} x_i = (\sum_{i \in I_0} x_i) + (\sum_{i \in I_1} x_i) = \sum_{i=0}^{1} (\sum_{j \in I_i} x_j)$ so $1 \in X$

2. Assume that $n \in X$ then if $\{x_i\}_{i \in I}$ is a family of elements with finite support is such that $I = \bigcup_{i \in \{0, \ldots, n+1\}} I_i$ then $I = \left(\bigcup_{i \in \{0, \ldots, n\}} I_i\right) \bigcup I_{n+1}$ (see 1.107) and we have then

$$\begin{aligned}
\sum_{i \in I} x_i &\underset{\text{previous lemma}}{=} \sum_{i \in \bigcup_{i \in \{0, \ldots, n\}} I_i} x_i + \sum_{i \in I_{n+1}} x_i \\
&\underset{n \in X}{=} \left(\sum_{i=0}^{n} \left(\sum_{j \in I_i} x_j\right)\right) + \sum_{i \in I_{n+1}} x_i \\
&= \sum_{i=0}^{n+1} \left(\sum_{j \in I_i} x_j\right)
\end{aligned}$$

and thus $n + 1 \in X$

Using 4.77 we have that $X = \mathbb{N}$ □

Theorem 10.47. *Let $\langle A, + \rangle$ be a abelian semi-group, $\{x_i\}_{i \in I}$ a family of elements in A with finite support and $\{I_i\}_{j \in J}$ a family such that $\forall i, j \in J$ with $j \neq i$ we have $I_i \bigcap I_j = \emptyset$ and $\bigcup_{j \in J} I_j = I$ then we have $\sum_{i \in I} x_i = \sum_{j \in J} \left(\sum_{i \in I_j} x_i\right)$*

Proof. Define

$$J_0 = \left\{ j \in J | I_j \bigcap \text{support}(\{x_i\}_{i\in I}) \neq \emptyset \right\} \subseteq J \tag{10.2}$$

then trivially $\bigcup_{j\in J_0} I_j \bigcap \text{support}(\{x_i\}_{i\in I}) \subseteq \text{support}(\{x_i\}_{i\in I})$, also if $i \in \text{support}(\{x_i\}_{i\in I}) \underset{I=\bigcup_{j\in J} I_j}{\Rightarrow} \exists j \in J$ with $i \in I_j \bigcap \text{support}(\{x_i\}_{i\in I}) \Rightarrow I_j \bigcap \text{support}(\{x_i\}_{i\in I}) \neq \emptyset \Rightarrow j \in J_0 \Rightarrow i \in \bigcup_{j\in J_0} I_j \bigcap \text{support}(\{x_i\}_{i\in I})$ giving that

$$\bigcup_{j\in J_0} I_j \bigcap \text{support}(\{x_i\}_{i\in I}) \;=\; \text{support}(\{x_i\}_{i\in I}) \tag{10.3}$$

Now as $\forall j \in J_0$ we have by definition that $I_j \bigcap \text{support}(\{x_i\}_{i\in I}) \neq \emptyset$ we can use the axiom of choice (see 2.201) to find a function $c: J_0 \rightarrow \bigcup_{j\in J_0} I_j \bigcap \text{support}(\{x_i\}_{i\in I}) = \text{support}(\{x_i\}_{i\in I})$ such that $\forall j \in J_0$ we have $c(j) \in I_j \bigcap \text{support}(\{x_i\}_{i\in I})$. Now if $j_1, j_2 \in J_0$ is such that $c(j_1) = c(j_2)$ then $c(j_1), c(j_2) \in I_{j_1} \bigcap \text{support}(\{x_i\}_{i\in I}) \bigcap I_{j_2} \bigcap \text{support}(\{x_i\}_{i\in I}) \Rightarrow c(j_1), c(j_2) \in I_{j_1} \bigcap I_{j_2} \underset{I_j's\,\text{are pairwise disjoint}}{=} c(j_1) = c(j_2)$ proving that c is a injection. So $J_0 \approx c(J_0) \subseteq \text{support}(\{x_i\}_{i\in I})$ which is finite and thus by 5.36 we have that

$$J_0\,\text{must be finite.} \tag{10.4}$$

As J_0 is finite there exists a $b: \{0, \ldots, n-1\} \rightarrow J_0$ which is a bijection. Also as $I_j \bigcap \text{support}(\{x_i\}_{i\in I}) \subseteq \text{support}(\{x_i\}_{i\in I})$ which is finite we have that $S_i = \sum_{i\in I_j \bigcap \text{support}(\{x_i\}_{i\in I})} x_i$ is well defined and we have also that $\sum_{j\in J_0} S_i$ is well defined. We have then $\sum_{j\in J_0} S_i = \sum_{j\in J_0} \left(\sum_{i\in I_j \bigcap \text{support}(\{x_i\}) i\in I} x_i\right) \underset{\text{definition}}{=} \sum_{j=0}^{n-1} \left(\sum_{i\in I_{b(j)} \bigcap \text{support}(\{x_i\}_{i\in I})} x_i\right) = \sum_{i\in \text{support}(\{x_i\}_{i\in I})} x_i$ (using 10.46, 10.2 and the fact that $I_{b(j)} \bigcap \text{support}(\{x_i\}_{i\in I})$ are disjoints (as b is a bijection). So we have proved that

$$\sum_{j\in J_0} \left(\sum_{i\in I_j \bigcap \text{support}(\{x_i\}_{i\in I})} x_i \right) = \sum_{i\in \text{support}(\{x_i\}_{i\in I})} x_i \tag{10.5}$$

Now if $k \in \text{support}\left(\left\{\sum_{j\in I_i \bigcap \text{support}(\{x_l\}_{l\in I})} x_j\right\}_{i\in J_0}\right)$ then $\sum_{j\in I_k \bigcap \text{support}(\{x_l\}_{l\in I})} x_i \neq 0 \Rightarrow \text{support}(\{x_l\}_{l\in I_k \bigcap \text{support}(\{x_l\}_{l\in I})}) \neq \emptyset \Rightarrow \exists l \in I_k \bigcap \text{support}(\{x_l\}_{l\in I}) \vdash x_l \neq 0 \Rightarrow I_k \bigcap \text{support}(\{x_l\}_{l\in I}) \neq \emptyset \Rightarrow k \in J_0$ so we have

$$\text{support}\left(\left\{ \sum_{j\in I_i \bigcap \text{support}(\{x_l\}_{l\in I})} x_j \right\}_{i\in J_0}\right) \subseteq J_0 \subseteq I \tag{10.6}$$

so using 10.42 and 10.5,10.6 twice we have

$$\sum_{j\in I} \left(\sum_{j\in I_i \bigcap \text{support}(\{x_l\}_{l\in I})} x_i \right) = \sum_{i\in I} x_i \tag{10.7}$$

also if $l \in \text{support}(\{x_i\}_{i\in I_k}) \Rightarrow l \in I_k \wedge x_l \neq 0 \Rightarrow l \in I_k \bigcap \text{support}(\{x_i\}_{i\in I}) \Rightarrow l \in I_k$ giving

$$\text{support}(\{x_i\}_{i\in I_k}) \subseteq I_k \bigcap \text{support}(\{x_i\}_{i\in I}) \subseteq I_k \tag{10.8}$$

so that we can use 10.7,10.8 and 10.42 to find finally

$$\sum_{j\in J}\left(\sum_{i\in I_j} x_i\right)=\sum_{i\in I} x_i \qquad \square$$

Corollary 10.48. *Let $\langle A,+\rangle$ be a abelian semi-group, I,J sets and $\{x_{i,j}\}_{(i,j)\in I\times J}$ a family of elements in A with finite support then $\sum_{(i,j)\in I\times J}x_{i,j}=\sum_{i\in I}(\sum_{j\in J}x_{i,j})=\sum_{j\in J}(\sum_{i\in I}x_{i,j})$*

Proof.

1. Define $\forall i\in I\;\; J_i=\{i\}\times J$ then we have clearly $I\times J=\bigcup_{i\in I}J_i$ and if $i,j\in I$ with $i\neq j$ we have if $(n,m)\in J_i\bigcap J_j\Rightarrow n=i,n=j\Rightarrow i=j$ contradicting $i\neq j$ so we have $J_i\bigcap J_j=\emptyset$. Using then the previous theorem 10.47 we have that $\sum_{(i,j)\in I\times J}x_{i,j}=\sum_{i\in I}\left(\sum_{(k,l)\in J_i}x_{k,l}\right)$. Define now the bijection $b_i\colon J\rightarrow J_i$ by $j\rightarrow b_i(j)=(i,j)$ then we have by 10.44 that $\sum_{(k,l)\in J_i}x_{k,l}=\sum_{l\in J}x_{b_i(k)}=\sum_{l\in J}x_{i,l}$ giving that $\sum_{(i,j)\in I\times J}=\sum_{i\in I}(\sum_{j\in J}x_{i,j})$.

2. Define $\forall j\in J\;\; I_j=I\times\{j\}$ then we have clearly $I\times J=\bigcup_{i\in j}I_i$ and if i, $j\in J$ with $i\neq j$we

3. have if $(n,m)\in I_i\bigcap I_j\Rightarrow m=i,m=j\Rightarrow i=j$ contradicting $i\neq j$ so we have $I_i\bigcap I_j=\emptyset$. Using then the previous theorem 10.47 we have that $\sum_{(i,j)\in I\times J}x_{i,j}=\sum_{i\in J}\left(\sum_{(k,l)\in I_i}x_{k,l}\right)$. Define now the bijection $h_i\colon I\rightarrow I_i$ by $j\rightarrow h_i(j)=(j,i)$ then we have by 10.44 that $\sum_{(k,l)\in J_i}x_{k,l}=\sum_{l\in J}x_{h_i(l)}=\sum_{l\in J}x_{l,i}$ giving that $\sum_{(i,j)\in I\times J}=\sum_{i\in J}(\sum_{j\in I}x_{j,i})$. $\square$

Theorem 10.49. *Let $\langle A,+\rangle$ be a abelian semi-group, $n\in\mathbb{N}\setminus\{1\}$, $\{N_i\}_{i\in\{1,\ldots,n\}}$ such that N_i is finite and non empty and $f\colon\prod_{i\in\{1,\ldots,n\}}A_i\rightarrow A$ then $\sum_{\gamma\in\prod_{i\in\{1,\ldots,n\}}A_i}f(\gamma)=\sum_{\gamma\in\prod_{i\in\{1,\ldots,n\}}A_i}f(\gamma_1,\ldots,\gamma_n)=\sum_{\gamma\in\prod_{i\in\{1,\ldots,k-1\}}}(\sum_{i\in A_n}f(\lambda_1,\ldots,\gamma_n,i))$*

Proof. Define $\{x_{(i,j)}\}_{(i,j)\in(\prod_{i\in\{1,\ldots,n-1\}}A_i)\times A_n}$ by $x_{(i,j)}=f(i_1,\ldots,i_{i-1},j)$ then we have

$$\begin{aligned}\sum_{(i,j)\in(\prod_{i\in\{1,\ldots,n-1\}}A_i)\times A_n}x_{(i,j)} &\underset{10.48}{=} \sum_{i\in\prod_{i\in\{1,\ldots,n-1\}}A_i}\left(\sum_{j\in A_n}x_{(i,j)}\right)\\ &= \sum_{i\in\prod_{i\in\{1,\ldots,n-1\}}A_i}\left(\sum_{j\in A_n}f(i_1,\ldots,i_{n-1},j)\right)\end{aligned}\tag{10.9}$$

Further

$$\begin{aligned}\sum_{(i,j)\in(\prod_{i\in\{1,\ldots,n-1\}}A_i)\times A_n} x_{(i,j)} &= \sum_{(i,j)\in(\prod_{i\in\{1,\ldots,n-1\}}A_i)\times A_n} x_{((i_1,\ldots,i_{n-1}),j)}\\ &\underset{5.102}{=} \sum_{(i_1,\ldots,i_n)\in\prod_{i\in\{1,\ldots,n\}}A_i} x_{(i_1,\ldots,i_n)}\\ &= \sum_{(i_1,\ldots,i_n)\in\prod_{i\in\{1,\ldots,n\}}A_i} f(i_1,\ldots,i_n)\end{aligned}$$

Using this with 10.9 proves the theorem. □

Lemma 10.50. *If $\langle A,+\rangle$ is a semi-group and $\{x_i\}_{i\in\{0,\ldots,n\}}$ is a set of elements in A such that $\forall i\in\{1,\ldots,0\}$ we have that $x_i=0$ (the neutral element of the semi-group) then $\sum_{i=0}^n x_i=0$*

Proof. This is proved by induction on n. So let $X=\{n\in\mathbb{N}_0|\text{if}\{x_i\}_{i\in\{0,\ldots,n\}}\text{is a family of elements in A that are all 0 then } \sum_{i=0}^n x_i\ \}$ then we have

1. If $n=0$ then $\{x_i\}_{i\in\{0,\ldots,0\}}$ has $x_0=0$ then $\sum_{i=0}^0 x_i=x_0=0$ so $0\in X$
2. If $n\in X$ then if $\{x_i\}_{i\in\{0,\ldots,n+1\}}$ of $0's$ we have $\sum_{i=0}^{n+1} x_i=(\sum_{i=0}^n x_i)+x_{n+1}\underset{x_{n+1}=0}{=}\sum_{i=0}^n x_i\underset{n\in X}{=}0$

So using mathematical induction we have $X=\mathbb{N}_0$ proving the theorem. □

Theorem 10.51. *If $\langle A,+\rangle$ is a abelian semi-group and $\{x_i\}_{i\in I}$ is a family of elements in A with I is finite and non-empty (and thus have finite support) then if $h: S_n\to I$ is a bijection (which must exists as I is finite) we have that $\sum_{i\in I}x_i=\sum_{i=0}^{n-1}x_{h(i)}$*

Proof. Let $K=\text{support}(\{x_i\}_{i\in I})$ then if $K=I$ we have if $h: S_n\to I=\text{support}(\{x_i\}_{i\in I})$ is a bijection that by definition of $\sum_{i\in I}x_i=\sum_{i=0}^{n-1}x_{h(i)}$. So assume that $K\subset I$ then if $m=\#(K)$ we have by 5.45 that $\#(I\backslash K)=n-m$. So there exists bijections $f_1: \{0,\ldots,m-1\}\to K$ and $f_2: \{0,\ldots,(n-m)-1\}\to I\backslash K$, define then the (easily proved) bijection $f_3: \{m,\ldots,n-1\}\to I\backslash K$: $i\to f_2(i-m)$. This allows us to create a bijection (see 2.43) $f_1\bigcup f_3: \{0,\ldots,n-1\}\to I$. Then $\sum_{i=0}^{n-1}x_{(f_1\bigcup f_3)(i)}\underset{10.26}{=}\sum_{i=0}^{m-1}x_{(f_1\bigcup f_3)(i)}+\sum_{i=m}^{n-1}x_{(f_1\bigcup f_3)(i)}=\sum_{i=0}^{n-1}x_{(f_1)(i)}+\sum_{i=m}^{n-1}x_{(f_3)(i)}\underset{K\,\text{is}\,\text{support}(\{x_i\}_{i\in I})}{=}\sum_{i\in I}x(i)+\sum_{i=m}^{n-1}x_{(f_3)(i)}=\sum_{i\in I}x_i+\sum_{i=0}^{(n-m)-1}x_{(f_3)(i+m)}\underset{\text{previous}\,\text{lemma}\,\text{and}\,x_{(f_3)((i+m))}=x_{(f_2(i))}=0}{=}\sum_{i\in I}x_i$ or summarized $\sum_{i=0}^{n-1}x_{(f_1\bigcup f_3)(i)}=\sum_{i\in I}x_i$. Now if $h: S_n=\{0,\ldots,n-1\}\to I$ is a bijection then $\sigma=(f_1\bigcup f_3)^{-1}\circ h: S_n\to S_n$ is a bijection (or permutation) so using 10.19 we have that $\sum_{i\in I}x_i=\sum_{i=0}^{n-1}x_{(f_1\bigcup f_3)(i)}=\sum_{i=0}^{n-1}x_{(f_1\bigcup f_3)_{((f_1\bigcup f_3^{-1})\circ h)(i)}}=\sum_{i=0}^{n-1}x((f_1\bigcup f_3)((f_1\bigcup f_3)^{-1}(h(i))))=\sum_{i=0}^{n-1}x(h(i))=\sum_{i=0}^{n-1}x_{h(i)}$ proving the theorem. □

A variant of the above theorem is expressed in the following lemma

Lemma 10.52. *If $\langle A,+\rangle$ is a abelian semi-group and $\{x_i\}_{i\in I}$ a family of elements in A with $\mathrm{support}(\{x_i\}_{i\in I})$ is finite and $\mathrm{support}(\{x_i\}_{i\in I})\subseteq J\subseteq I$ where J is a finite set then if $h: S_n\rightarrow J$ is a bijection we have that $\sum_{i=0}^{n-1} x_{h(i)}=\sum_{i\in I} x_i$*

Proof. As $\mathrm{support}(\{x_i\}_{i\in I})=\mathrm{support}(\{x_i\}_{i\in J})$ we have if $g: S_n\rightarrow \mathrm{support}(\{x_i\}_{i\in I})=\mathrm{support}(\{x_i\}_{i\in J})$ is bijection that $\sum_{i=0}^{n-1} x_i \underset{10.51}{=} \sum_{i\in J} x_i=\sum_{i\in I} x_i$ □

Theorem 10.53. *If $\langle A,+\rangle$ is a abelian semi-group and $\{x_i\}_{i\in I}$ and $\{y_i\}_{i\in I}$ are family of sets in A with finite support then $\{x_i+y_i\}_{i\in I}$ is a family with finite support and $\sum_{i\in I}(x_i+y_i)=(\sum_{i\in I} x_i)+(\sum_{i\in} y_i)$*

Proof. If $i\in\mathrm{support}(\{x_i+y_i\}_{i\in I})$ then $x_i+y_i\neq 0$, now if $x_i=0=y_i\Rightarrow x_i+y_i=0$ a contradiction so we have either $x_i\neq 0$ or $y_i\neq 0$ and thus $i\in\mathrm{support}(\{x_i\}_{i\in I})\bigcup \mathrm{support}(\{y_i\}_{i\in I})$ and thus $\mathrm{support}(\{x_i+y_i\}_{i\in I})\subseteq \mathrm{support}(\{x_i\}_{i\in I})\bigcup \mathrm{support}(\{y_i\}_{i\in I})$. As $\mathrm{support}(\{x_i\}_{i\in I}), \mathrm{support}(\{y_i\}_{i\in I})$ are finite we have $\mathrm{support}(\{x_i\}_{i\in I})\bigcup \mathrm{support}(\{y_i\}_{i\in I})$ is finite (see 5.37) and thus by 5.36 we have that $\mathrm{support}(\{x_i+y_i\}_{i\in I})$ is finite. If $h: S_n\rightarrow \mathrm{support}(\{x_i\}_{i\in I})\bigcup \mathrm{support}(\{y_i\}_{i\in I})$ is a bijection then as $\mathrm{support}(\{x_i+y_i\}_{i\in I})\subseteq \mathrm{support}(\{x_i\}_{i\in I})\bigcup \mathrm{support}(\{y_i\}_{i\in I})\subseteq I$ and $\mathrm{support}(\{x_i\}_{i\in I}), \mathrm{support}(\{y_i\}_{i\in I})\subseteq \mathrm{support}(\{x_i\}_{i\in I})\bigcup \mathrm{support}(\{y_i\}_{i\in I})$ we have $\sum_{i\in I}(x_i+y_i) \underset{\text{previous lemma}}{=} \sum_{i=0}^{n-1}(x_{h_i}+y_{h_i}) \underset{10.11}{=} (\sum_{i=0}^{n-1} x_{h_i})+(\sum_{i=0}^{n-1} y_{h_i}) \underset{\text{previous lemma}}{=} \sum_{i\in I} x_i+\sum_{i\in I} y_i$ □

Corollary 10.54. *If $\langle A,+\rangle$ is a abelian additive semi-group with unit 0 and $\{x_i\}_{i\in I}$ is a finite family of elements in A then given $a\in A$ we have for $\{x_i+a\}_{i\in I}$ that $\sum_{i\in I}(x_i+a)=(\sum_{i\in I} x_i)+(a\cdot\#(I))$*

Proof. Define $\{a_i\}_{i\in I}$ by $\forall i\in I$ we have $a_i=a$ then $\{x_i+a\}_{i\in I}=\{x_i+a_i\}_{i\in I}$, also let $h: S_n\rightarrow I$ be a bijection where $n=\#(I)$ then $\sum_{i\in I}(x_i+a)=\sum_{i\in I}(x_i+a_i)=\sum_{i\in I} x_i+\sum_{i\in I} a_i=\sum_{i\in I} x_i+\sum_{i=0}^{n-1} a_{h_i} \underset{a_{h_i}=a \text{ and } 10.6}{=} \sum_{i\in I} x_i+a\cdot n=\sum_{i\in I} x_i+a\cdot\#(I)$ □

Corollary 10.55. *If $\langle A,\cdot\rangle$ is a abelian multiplicative semi-group with unit 1 and $\{x_i\}_{i\in I}$ is a finite family of elements in A then given $a\in A$ we have for $\{x_i\cdot a\}_{i\in I}$ that $\prod_{i\in I}(x_i\cdot a)=(\prod_{i\in I} x_i)\times(a^{\#(I)})$*

Proof. Define $\{a_i\}_{i\in I}$ by $\forall i\in I$ we have $a_i=a$ then $\{x_i\cdot a\}_{i\in I}=\{x_i\cdot a_i\}_{i\in I}$, also let $h: S_n\rightarrow I$ be a bijection where $n=\#(I)$ then $\prod_{i\in I}(x_i\cdot a)=\prod_{i\in I}(x_i\cdot a_i)=\prod_{i\in I} x_i\cdot\prod_{i\in I} a_i=\prod_{i\in I} x_i\cdot\prod_{i=0}^{n-1} a_{h_i} \underset{a_{h_i}=a \text{ and } 10.7}{=} \prod_{i\in I} x_i\cdot a^n=\sum_{i\in I} x_i\cdot a^n$ □

If we have a ring then we have the following interesting properties for

Theorem 10.56. *If $\langle R,+,\cdot\rangle$ is a commutative ring with neutral element and unit element 1 then if $\{x_i\}_{i\in I}$ is a family of elements in A with finite support then*

1. *If $\exists i\in I$ with $x_i=0$ then $\prod_{i\in I}x_i=0$*
2. *If $\langle R,+,\cdot\rangle$ is a integral domain (such that if $x,y\in R\setminus\{0\}$ then $x\cdot y\neq 0$) and $\forall i\in I$ we have that $x_i\neq 0$ then* $\text{support}(\{x_i^{-1}\}_{i\in I})=\text{support}(\{x_i\}_{i\in I})$ *is finite (so that $\prod_{i\in I}x_i^{-1}$ is defined), $\prod_{i\in I}x_i\neq 0$ and $(\prod_{i\in I}x_i)^{-1}=\prod_{i\in I}x^{-1}$*

Proof.

1. $\exists i\in I$ such that $x_i=0$ then as I is the disjoint union of $\{i\}$ and $I\setminus\{i\}$ we have that $\prod_{i\in I}x_i=\left(\prod_{i\in I\setminus\{i\}}x_i\right)\cdot\left(\prod_{i\in\{i\}}x_i\right)=(\prod_{i\in I}x_i)\cdot x_i=(\prod_{i\in I}x_i)\cdot 0=0$
2. First if $i\in\text{support}(\{x_i\}_{i\in I})$ then $x_i\neq 1$, if now $x_i^{-1}=1$ then $1=x_1\cdot x_i^{-1}=x_i 1=x_i\neq 1$ yielding contradiction so that $x_i^{-1}\neq 1\Rightarrow i\in\text{support}(\{x_i^{-1}\}_{i\in I})$. If $i\in\text{support}(\{x_i^{-1}\}_{i\in I})$ then $x_i^{-1}\neq 1$,if now $x_i=1$ then $1=x_i^{-1}\cdot x_i=x_i^{-1}\cdot 1=x_i^{-1}\neq 1$,a contradiction so that $x_i\neq 1$ and thus $i\in\text{support}(\{x_i\}_{i\in I})$. We prove now by induction that if $\{x_i\}_{i\in\{0,\ldots,n\}}$ is such that $\forall i\in\{1,\ldots,n\}$ we have $x_i\neq 0$ then $\prod_{i=0}^n x_i^{-1}=(\prod_{i=0}^n x_i)^{-1}$ so let $S=\{n\in\mathbb{N}|\text{if }\{x_i\}_{i\in\{0,\ldots,n\}}\text{ is such that }\forall i\in\{0,\ldots,n\}\text{ we have }x_i\neq 0\text{ then}\prod_{i=0}^n x_i\neq 0\text{ and }\prod_{i=0}^n x_i^{-1}=(\prod_{i=0}^n x_i)^{-1}\}$ then we have

   a. If $n=0$ then $0\neq x_0=\prod_{i=0}^0 x_i^{-1}=x_0^{-1}=(\prod_{i=0}^0 x_i)^{-1}$ so that $0\in S$

   b. If $n\in S$ then $\prod_{i=0}^{n+1}x_i=(\prod_{i=0}^n x_i)\cdot x_{n+1}\neq 0$ (as $n\in S$ and $x_{n+1}\neq 0$), also $\prod_{i=0}^{n+1}x_i^{-1}=(\prod_{i=0}^n x_i^{-1})\cdot x_{n+1}^{-1}\underset{n\in S}{=}(\prod_{i=0}^n x_i)^{-1}\cdot x_{n+1}^{-1}=((\prod_{i=0}^n x_i)\cdot x_{n+1})^{-1}=(\prod_{i=0}^{n+1}x_i)^{-1}$ so that $n+1\in S$.

   Now if $\text{support}(\{x_i^{-1}\}_{i\in I})=\text{support}(\{x_i\}_{i\in I})=\emptyset$ then $\prod_{i\in I}x_i^{-1}=1=1^{-1}=(\prod_{i\in I}x_i)^{-1}$, if $\text{support}(\{x_i\}_{i\in I})=\text{support}(\{x_i^{-1}\}_{i\in I})\neq\emptyset$ then there exists a bijection b: $\{0,\ldots,n\}\to\text{support}(\{x_i\}_{i\in I})=\text{support}(\{x_i^{-1}\}_{i\in I})$ such that $\prod_{i\in I}x_i^{-1}=\prod_{i=0}^n x_{b(i)}^{-1}=(\prod_{i=0}^n x_{b(i)})^{-1}=(\prod_{i\in I}x_i)^{-1}$ □

Lemma 10.57. *If $\{\langle A_i,+_i\rangle\}_{i\in I}$ is a family of abelian semi-groups then by 3.15 we have that $\langle\prod_{i\in I}A_i,+\rangle$ is a abelian semi-group. If now $\{x_j\}_{j\in J}$ is a family of elements in $\prod_{i\in I}A_i$ with finite support then $\forall i\in I$ we have $(\sum_{j\in J}x_j)(i)\underset{\text{notation}}{=}(\sum_{j\in J}x_j)_i=\sum_{j\in J}(x_j)_i\underset{\text{notation}}{=}\sum_{j\in J}(x_j)(i)$*

Proof. As $\{x_j\}_{j\in J}$ has finite support then we have the following cases to consider:

1. **($\text{support}(\{x_j\}_{j\in J})=\emptyset$)** Then if $i\in I$ and $j\in J\Rightarrow x_j=0\Rightarrow(x_j)(i)=0\Rightarrow$ $\text{support}(\{(x_j)(i)\})=\emptyset$ and we have $(\sum_{j\in J}x_j)(i)=0(i)=0=\sum_{j\in J}(x_j)(i)$

2. $\mathbf{(support(\{x_j\}_{j\in J})\neq\emptyset)}$ then $\forall i\in I$ if $j\in\text{support}(\{x_j(i)\})\Rightarrow x_j(i)\neq 0\Rightarrow x_j\neq 0\Rightarrow j\in\text{support}(\{x_j\}_{j\in J})$ giving $\text{support}(\{x_j(i)\}_{j\in J})\subseteq\text{support}(\{x_j\}_{j\in J})$ proving that $\text{support}(\{x_j(i)\}_{j\in J})$ is finite so the sums are defined and we have a bijection $b:\{0,\ldots,n-1\}\rightarrow\text{support}(\{x_j\}_{j\in J})$ then by 10.52 we have $(\sum_{j\in J}x_j)(i)=(\sum_{j=0}^{n-1}x_{b_j})(i)\underset{10.10}{=}\sum_{j=0}^{n-1}(x_{b_j})(i)\underset{10.52}{=}\sum_{j\in J}x_j(i)$ □

Theorem 10.58. *Let $\mathbb{R}$ (or $\mathbb{R}_{\mathbb{C}}$) be the real field (or the real field embedded in $\mathbb{C}$) and $\{x_i\}_{i\in I}$ be such that $\forall i\in\sup(\{x_i\}_{i\in I})$ we have $x_i=c$ then we have that $\sum_{i\in I}x_i=(\#(\sup(\{x_i\}_{i\in I})))\cdot c$*

Proof. If $b:\{0,\ldots,n-1\}\rightarrow\sup(\{x_i\}_{i\in I})$ is a bijection where $n=\#(\sup(\{x_i\}_{i\in I}))$ then we have that $\sum_{i\in I}x_i=\sum_{i=0}^{n-1}x_{b(i)}=\sum_{i=0}^{n-1}c\underset{10.31}{=}((n-1)+1)\cdot c=n\cdot c=\#(\sup(\{x_i)_{i\in I}))$ □

Theorem 10.59. *Let $\langle F,+,\cdot\rangle$ be a field $\alpha\in F$ and $\{x_i\}_{i\in I}$ a family with finite support then $\sum_{i\in I}\alpha\cdot x_i=\alpha\cdot\sum_{i\in I}x_i$*

Proof. Let $b:\{0,\ldots,n-1\}\rightarrow\text{support}(\{x_i\}_{i\in I})$ then we have $\sum_{i\in I}\alpha\cdot x_i=\sum_{i=0}^{n}\alpha\cdot x_{b(i)}=\alpha\cdot\sum_{i=0}^{n}x_{b(i)}=\alpha\cdot\sum_{i\in I}x_i$ □

Corollary 10.60. *Let $\langle A,+,\cdot\rangle$ be a field, I,J sets and $\{x_i\}_{i\in I}$, $\{y_i\}_{i\in I}$ families of elements in I, J with finite support then $\sum_{(i,j)\in I\times J}(x_i\cdot x_j)=(\sum_{i\in I}x_i)\cdot(\sum_{i\in J}x_j)$*

Proof. As for a field $\langle A,+\rangle$ is a Abelian group we have by 10.48 that $\sum_{(i,j)\in I\times J}(x_i\cdot y_j)=\sum_{i\in I}(\sum_{j\in J}(x_i\cdot y_j))\underset{10.59}{=}\sum_{i\in I}x_i\cdot(\sum_{j\in J}x_j)\underset{10.59}{=}(\sum_{i\in I}x_i)\cdot(\sum_{j\in J}x_j)$ □

Theorem 10.61. *Let $\langle R,+,\cdot\rangle$ be a ring, $n\in\mathbb{N}$ and $\{\{a_{i,j}\}_{j\in N_i}\}_{i\in\{1,\ldots,n\}}$ such that $\forall i\in\{1,\ldots,n\}$ N_i is finite and non empty then $\prod_{i\in\{1,\ldots,n\}}(\sum_{i\in N_i}a_{i,j})=\sum_{\gamma\in\prod_{i\in\{1,\ldots,n\}}N_i}(\prod_{i\in\{1,\ldots,n\}}a_{i,\gamma_i})$*

Proof. We prove this by induction on n so let $\mathcal{S}=\left\{n\in\mathbb{N}|\forall\{\{a_{i,j}\}_{j\in N_i}\}_{i\in\{1,\ldots,n\}},\right.$ such that N_i is finite non empty, we have $\left.\prod_{i\in\{1,\ldots,n\}}(\sum_{i\in N_i}a_{i,j})=\sum_{\gamma\in\prod_{i\in\{1,\ldots,n\}}N_i}(\prod_{i\in\{1,\ldots,n\}}a_{i,\gamma_i})\right\}$ then

$\mathbf{1\in\mathcal{S}}$. If $\{\{a_{i,j}\}_{j\in N_i}\}_{i\in\{1,\ldots,1\}}$ N_1 finite non empty then

$$\begin{aligned}\prod_{i\in\{1,\ldots,1\}}\left(\sum_{j\in N_i}a_{i,j}\right)&=\sum_{j\in N_1}a_{1,j}\\&\underset{5.102}{=}\sum_{\gamma\in\prod_{i\in\{1,\ldots,1\}}N_i}a_{1,\gamma_1}\\&=\prod_{i\in\{1,\ldots,1\}}\left(\sum_{\gamma\in\prod_{i\in\{1,\ldots,1\}}N_i}a_{i,\gamma_i}\right)\end{aligned}$$

proving that $1\in\mathcal{S}$

$\boldsymbol{n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}}$. Let $\{\{a_{i,j}\}_{j \in N_i}\}_{i \in \{1,\ldots,n+1\}}$ be a finite family such that N_i is finite and non empty then

$$\begin{aligned}
\prod_{i \in \{1,\ldots,n+1\}} \left(\sum_{j \in A_i} a_{i,j}\right) &= \left(\prod_{i \in \{1,\ldots,n\}} \left(\sum_{j \in A_i} a_{i,j}\right)\right) \cdot \left(\sum_{j \in A_{n+1}} a_{n+1,j}\right) \\
&\underset{n \in \mathcal{S}}{=} \left(\sum_{\gamma \in \prod_{i \in \{1,\ldots,n\}} N_i} \left(\prod_{i \in \{1,\ldots,n\}} a_{i,\gamma_i}\right)\right) \cdot \\
&\quad \left(\sum_{j \in A_{n+1}} a_{n+1,j}\right) \\
&\underset{10.59}{=} \sum_{\gamma \in \prod_{i \in \{1,\ldots,n\}} N_i} \left(\left(\prod_{i \in \{1,\ldots,n\}} a_{i,\gamma_i}\right) \cdot \right. \\
&\quad \left. \sum_{j \in A_{n+1}} a_{n+1,j}\right) \\
&\underset{10.59}{=} \sum_{\gamma \in \prod_{i \in \{1,\ldots,n\}} N_i} \left(\sum_{j \in A_{n+1}} \left(\left(\prod_{i \in \{1,\ldots,n\}} a_{i,\gamma_i}\right) \cdot \right.\right. \\
&\quad \left.\left. a_{n+1,j}\right)\right) \\
&\underset{10.49}{=} \sum_{\gamma \in \prod_{i \in \{1,\ldots,n+1\}} A_i} \left(\left(\prod_{i \in \{1,\ldots,n\}} a_{i,\gamma_i}\right) \cdot \right. \\
&\quad \left. a_{n+1,\gamma_{n+1}}\right) \\
&= \sum_{\gamma \in \prod_{i \in \{1,\ldots,n+1\}} A_i} \left(\prod_{i \in \{1,\ldots,n+1\}} a_{i,\gamma_i}\right)
\end{aligned}$$

proving that $n+1 \in \mathcal{S}$ □

Theorem 10.62. *Let* $\mathbb{R}$ *(or* $\mathbb{R}_{\mathbb{C}}$*) be the real field (embedded real field in the complex numbers,* $\{x_i\}_{i \in I}$ *a family in* $\mathbb{R}$ *(or* $\mathbb{R}_{\mathbb{C}}$*) with finite support such that* $\forall i \in I$ *we have* $0 \leqslant x_i$ *then* $0 \leqslant (<) \sum_{i \in I} x_i$ *(where we have* $<$ *if* $\operatorname{support}(\{x_i\}_{i \in I}) \neq 0$*) and* $\sum_{i \in I} x_i = 0$ *if* $\operatorname{support}(\{x_i\}_{i \in I}) = 0$

Proof. We have the following cases to consider:

1. $\boldsymbol{(\operatorname{support}(\{x_i\}_{i \in I}) = \emptyset)}$ then $\sum_{i \in I} x_i = 0$ by definition

2. $\boldsymbol{(\operatorname{support}(\{x_i\}_{i \in I}) \neq \emptyset)}$ then let $b: \{0, \ldots, n-1\} \rightarrow \operatorname{support}(\{x_i\}_{i \in I})$ then we have that $n > 0 \Rightarrow x_{b(0)} > 0$ so that we have $\sum_{i \in I} x_i = \sum_{i=0}^{n-1} x_{b(i)} > 0$ using 10.32 □

Theorem 10.63. *Let* $\mathbb{R}$ *(or* $\mathbb{R}_{\mathbb{C}}$*) be the real field (embedded real field in the complex numbers,* $\{x_i\}_{i\in I}$*,* $\{y_i\}_{i\in I}$ *families in* $\mathbb{R}$ *(or* $\mathbb{R}_{\mathbb{C}}$*) such that* $\forall i\in I$ *we have* $x_i\leqslant y_i$ *then we have* $\sum_{i\in I}x_i\leqslant(<)\sum_{i\in I}y_i$ *(where we have* $<$ *if there is a* $x_i<y_i$*)*

Proof. Consider the family $\{y_i+(-x_i)\}_{i\in I}$ then we have $\forall i\in I$ that $y_i+(-x_i)\geqslant 0$ and we have thus by 10.62 that

$$\sum_{i\in I}(y_i+(-x_i))\geqslant 0(>0)\text{ (where we have }>\text{ if there exists a }y_i>x_i) \tag{10.10}$$

Now $\sum_{i\in I}(y_i+(-x_i))\underset{10.53}{=}\sum_{i\in I}y_i+\sum_{i\in I}(-x_i)\underset{10.59}{=}\sum_{i\in I}y_i+(-\sum_{i\in I}x_i)=\sum_{i\in I}y_i-\sum_{i\in I}x_i$ which together with 10.10 gives that $\sum_{i\in I}y_i\geqslant(>)\sum_{i\in I}x_i$ where we have $>$ if there exists a $y_i>x_i$ □

Theorem 10.64. *Let* $\delta\in\mathbb{R}_+$*,* $\{x_i\}_{i\in I}$ *is a family in* $\mathbb{R}_+$ *with finite* $\text{support}(\{x_i\}_{i\in I})=\{i\in I|x_i\neq 1\}\neq\emptyset$ *such that* $\forall i\in I$ *we have* $0<x_i<\delta$ *then* $0<\prod_{i\in I}x_i<\delta^{\#(\text{support}(\{x_i\}_{i\in I}))}<\delta<1$

Proof. First we prove by induction that $\prod_{i=0}^n x_i<\delta^{n+1}$ so let $S=\{n\in\mathbb{N}_0|\text{if }\{x_i\}_{i\in\{0,\ldots,n\}}\text{ is such that }\forall i\in\{0,\ldots,n\}\ 0<x_i<\delta\text{ then }0<\prod_{i=0}^n x_i<\delta^{n+1}\}$ then we have

1. If $n=0$ then $0<x_0=\prod_{i=0}^0 x_i=x_0<\delta=\delta^{0+1}$ so that $0\in S$
2. Let $n\in S$ and $\{x_i\}_{i\in(0,\ldots,n+1)}$ such that $\forall i\in\{0,\ldots,n+1\}$ we have $0<x_i<\delta$ then we have as $0<x_{n+1}<\delta`<1$ and $n\in S$ we have $0<(\prod_{i=0}^n x_i)\cdot x_{n+1}=\prod_{i=0}^{n+1}x_i=(\prod_{i=0}^n x_i)\cdot x_{n+1}<\delta^n\cdot x_{n+1}<\delta^n\cdot\delta<\delta^{n+1}$ so that $n+1\in S$

Second if $\emptyset\neq\text{support}(\{x_i\}_{i\in I})$ is finite then there exists a bijection b: $\{0,\ldots,\#(\text{support}(\{x_i\}_{i\in I}))-1\}\rightarrow\text{support}(\{x_i\}_{i\in I})$ and $0<\prod_{i=0}^{\#(\text{support}(\{x_i\}_{i\in I}))-1}x_{b(i)}=\prod_{i\in I}x_i=\prod_{i=0}^{\#(\text{support}(\{x_i\}_{i\in I}))-1}x_{b(i)}<\delta^{\#(\text{support}(x_i)_{i\in I})}$. The rest follows from 9.41 (13). □

10.2 Permutations

Definition 10.65. *Given* $n\in\mathbb{N}$ *we define* $P_n=\{\sigma|\sigma\colon\{1,\ldots,n\}\rightarrow\{1,\ldots,n\}\text{ is a permutation}\}$ *[*P_n *is the collection of all permutations on* $\{1,\ldots,n\}$ *[or equivalently the collection of bijections from* $\{1,\ldots,n\}$ *to* $\{1,\ldots,n\}$*]*

Theorem 10.66. $\langle P_n,\circ\rangle$ *forms a group (*$\circ$ *is composition of functions) called the permutation group*

Proof.

1. Given $\sigma,\tau\in P_n$ we have that $\sigma\circ\tau$ is again a element of σ (composition of bijections is again a bijection), so $\circ$ defines a function $P_n\times P_n\rightarrow P_n$

2. **(Associativity)** $\forall \sigma, \tau, \kappa \in P_n$ and $\forall i \in \{1, \ldots, n\}$ we have

$$\begin{aligned} ((\sigma \circ \tau) \circ \kappa)(i) &= (\sigma \circ \tau)(\kappa(i)) \\ &= \sigma(\tau(\kappa(i))) \\ &= \sigma((\tau \circ \kappa)(i)) \\ &= (\sigma \circ (\tau \circ \kappa))(i) \end{aligned}$$

   and thus we have $((\sigma \circ \tau) \circ \kappa) = \sigma \circ (\tau \circ \kappa)$

3. **(neutral element)** $\forall \sigma \in P_n$ we have $\sigma \circ 1_{\{1,\ldots,n\}} = \sigma = 1_{\{1,\ldots,n\}} \circ \sigma$ so $1_{\{1,\ldots,n\}}$ the identity mapping is the neutral element in P_n

4. **(inverse)** $\forall \sigma \in P_n$ we have as σ is a bijection the existence of a inverse σ^{-1} so that $\sigma \circ \sigma^{-1} = i_{\{1,\ldots,n\}} = \sigma^{-1} \circ \sigma$ □

Note that $\langle P_n, \circ \rangle$ is not commutative.

Example 10.67. Let $n \in \mathbb{N}$ then $\iota_n: \{1, \ldots, n\} \rightarrow \{1, \ldots, n\}$ defined by $j \rightarrow \iota_n(j) = n - j + 1$ is a permutation (further it is trivial that if $m \in \{1, \ldots, n\}$ then $(\iota_n)_{|\{1,\ldots,m\}} = \iota_m$). Here $\iota_n(n) = n - n + 1 = 1, \iota_n(n-1) = 2, \ldots, \iota_n(1) = n - 1 + 1 = n$.

Proof.

1. **(injectivity)** If $\iota_n(j) = \iota_n(k) \Rightarrow n - j = b - k \Rightarrow -j = -k \Rightarrow j = k$
2. **(surjectivity)** If $k \in \{1, \ldots, n\}$ take then $j = n - k + 1 \in \{1, \ldots, n\}$ so that $\iota_n(n - k + 1) = n - (n - k + 1) + 1 = k$ □

The following theorem shows how given a $\sigma \in P_n$ we construct a new $\sigma^{[i]} \in P_{n+1}$ such that $\sigma^{[i]}(n+1) = i$

Theorem 10.68. *Given $n \in \mathbb{N}$, $\sigma \in P_n$ and $i \in \{1, \ldots, n+1\}$ define then $\sigma^{[i]}: \{1, \ldots, n+1\} \rightarrow \{1, \ldots, n+1\}$ as follows:*

1. *If $i = n+1$ then $\forall j \in \{1, \ldots, n+1\}$ we have*

$$\begin{aligned} \sigma^{[i]}(j) &= \sigma(j) \text{ if } j \in \{1, \ldots, n\} \\ &= n+1 \text{ if } j = n+1 \end{aligned}$$

2. *If $i \in \{1, \ldots, n\}$ take then $k = \sigma^{-1}(i)$ (so that $\sigma(k) = i$) then $\forall j \in \{1, \ldots, n+1\}$ we have*

$$\begin{aligned} \sigma^{[i]}(j) &= \sigma(j) \text{ if } j \in \{1, \ldots, n\} \backslash \{k\} \\ &= i \text{ if } j = n+1 \\ &= n+1 \text{ if } j = k \end{aligned}$$

we have then that $\sigma^{[i]}$ is a bijection and thus $\sigma^{[i]} \in P_{n+1}$. Further from the definition of $\sigma^{[i]}$ it is immediate clear that $\sigma^{[i]}(n+1) = i$.

Proof. We have the following cases to consider

1. **($\boldsymbol{i = n+1}$)** If we define the bijection $\sigma': \{n+1\} \rightarrow \{n+1\}$ by $\sigma'(n+1) = n+1$ then as $\{1, \ldots, n\} \bigcap \{n+1\} = \emptyset$ and $\sigma^{[i]} = \sigma \bigcup \sigma'$ we have by 2.43 that $\sigma^{[i]}$ is a bijection.

2. **($i \in \{1, \ldots, n\}$)** If $k = \sigma^{-1}(i) \in \{1, \ldots, n\}$ we have using 2.51 that $\sigma_{|\{1, \ldots, n\} \backslash \{k\}}$: $\{1, \ldots, n\} \backslash \{k\} \rightarrow \{1, \ldots, n\} \backslash \{i\}$ is a bijection. And if we take then the bijection [because $n+1 \neq k$] σ': $\{k, n+1\} \rightarrow \{i, n+1\}$ defined by

$$\begin{array}{rcl} \sigma'(k) & = & n+1 \\ \sigma'(n+1) & = & i \end{array}$$

We have then trivially that $\sigma^{[i]} = \sigma_{|\{1, \ldots, n\} \backslash \{k\}} \bigcup \sigma'$. Using the fact that $\{1, \ldots, n\} \backslash \{k\} \bigcap \{n+1, k\} = \emptyset$ and $\{1, \ldots, n\} \backslash \{i\} \bigcap \{i, n+1\} = \emptyset$ we can use 2.43 to conclude that $\sigma^{[i]}$ is a bijection. □

We prove now that $\{1, \ldots, n+1\} \times P_n \approx P_{n+1}$

Theorem 10.69. *Let $n \in \mathbb{N}$ and Δ: $\{1, \ldots, n+1\} \times P_n \rightarrow P_{n+1}$ by $\Delta(i, \sigma) = \sigma^{[i]}$ (previous theorem guaranties that $\sigma^{[i]} \in P_{n+1}$). Δ is then proven to be a bijection.*

Proof.

1. **(injective)** Assume that $\Delta(k_1, \sigma_1) = \Delta(k_2, \sigma_2)$ then we have $\sigma_1^{[k_1]} = \sigma_2^{[k_2]}$. So $k_1 = \sigma_1^{[k_1]}(n+1) = \sigma_2^{[k_2]}(n+1) = k_2 \Rightarrow k_1 = k_2$. We can now consider two cases
   a. **($\boldsymbol{k_1 = k_2 = n+1}$)** If $j \in \{1, \ldots, n\}$ then $\sigma_1(j) = \sigma_1^{[n+1]}(j) = \sigma_1^{[k]}(j) = \sigma_2^{[k_2]}(j) = \sigma_2^{[n+1]}(j) = \sigma_2(j) \Rightarrow \sigma_1(j) = \sigma_2(j) \Rightarrow \sigma_1 = \sigma_2$
   b. **($\boldsymbol{k_1 = k_2 \neq n+1}$)** take now $l_1 = \sigma_1^{-1}(k_1), l_2 = \sigma_2^{-1}(k_2)$ then $\sigma_2^{[k_2]}(l_1) = \sigma_1^{[k_1]}(l_1) = n+1 = \sigma_2^{[k_2]}(l_2) \underset{\sigma_2^{[k_2]} \text{is a bijection}}{\Rightarrow} l_1 = l_2$. Now if $j \in \{1, \ldots, n\}$ then we have the following cases
      i. **($\boldsymbol{j = l_1 = l_2}$)** then $\sigma_1(j) = \sigma_1(l_1) = k_1 = k_2 = \sigma_2(l_2) = \sigma_2(j)$
      ii. **($\boldsymbol{j \neq l_1 = l_2}$)** then $\sigma_1(j) = \sigma_1^{[k_1]}(j) = \sigma_2^{[k_2]}(j) = \sigma_2(j)$

      so we conclude that $\sigma_1 = \sigma_2$ again.

   from (a) and (b) it follows then $\sigma_1 = \sigma_2$ proving injectivity.
2. **(surjective)** Let $\sigma \in P_{n+1}$ then we have two cases
   a. **($\boldsymbol{\sigma(n+1) = n+1}$)** then clearly $\sigma = (\sigma_{|\{1, \ldots, n\}})^{[n+1]} = \Delta(n+1, \sigma_{|\{1, \ldots, n\}})$
   b. **($\boldsymbol{\sigma(n+1) \in \{1, \ldots, n\}}$)** let now $k = \sigma(n+1) \in \{1, \ldots, n\}$ and $l = \sigma^{-1}(n+1)$ then $l \in \{1, \ldots, n\}$ [f $l = n+1 \Rightarrow \sigma(l) \in \{1, \ldots, n\} \Rightarrow \sigma(l) \neq n+1$ a contradiction so $l \in \{1, \ldots, n\}$] and if $j \in \{1, \ldots, n\} \backslash \{l\}$ we have $\sigma(j) \in \{1, \ldots, n\}$ [if $\sigma(j) = n+1 = \sigma(l) \underset{\sigma \text{ is bijective}}{\Rightarrow} l = j$ a contradicting $j \in \{1, \ldots, n\} \backslash \{l\}$]. Define now σ': $\{1, \ldots, n\} \rightarrow \{1, \ldots n\}$ by

$$\begin{array}{rcl} \sigma'(j) & = & \sigma(j) \text{ if } j \in \{1, \ldots, n\} \backslash \{l\} \\ & = & k \text{ if } j = l \end{array}$$

   We prove now that σ' is a bijection and thus $\sigma' \in P_n$.

Proof.

i. **(surjective)** Let $j \in \{1, \ldots, n\}$ then we must consider the following cases

   A. $(\boldsymbol{j = k})$ then $\sigma'(l) = k = j \Rightarrow \sigma'(l) = j$

   B. $(\boldsymbol{j \in \{1, \ldots, n\} \setminus \{k\}})$ by surjectivity of σ there exists a $i \in \{1, \ldots, n+1\}$ such that $\sigma(i) = j$. We can not have $i = n+1$ as then $j = \sigma(i) = \sigma(n+1) = k$ contradicting $j \in \{1, \ldots, n\} \setminus \{k\}$. Also we cannot have $i = l$ as then $j = \sigma(i) = \sigma(l) = n+1$ contradicting $j \in \{1, \ldots, n\} \setminus \{k\}$. So we must have $i \in \{1, \ldots, n\} \setminus \{l\}$ and thus $\sigma'(i) = \sigma(i) = j \Rightarrow \sigma'(i) = j$

   this proves surjectivity.

ii. **(injective)** Let $i, j \in \{1, \ldots, n\}$ be such that $\sigma'(i) = \sigma'(j)$ then we must consider the following cases

   A. $(\boldsymbol{\sigma'(i) = \sigma'(j) = k})$ then $i = l$ [if $i \neq l$ we have $i \in \{1, \ldots, n\} \setminus \{l\} \Rightarrow \sigma(n+1) = k = \sigma'(i) = \sigma(i) \underset{\sigma \text{ is a bijection}}{\Rightarrow} i = n+1$ contradicting $i \in \{1, \ldots, n\} \setminus \{l\}$], also $j = l$ [if $j \neq l$ we have $j \in \{1, \ldots, n\} \setminus \{l\} \Rightarrow \sigma(n+1) = k = \sigma'(j) = \sigma(j) \underset{\sigma \text{ is a bijection}}{\Rightarrow} j = n+1$ contradicting $j \in \{1, \ldots, n\} \setminus \{l\}$] so we conclude that $i = l = j \Rightarrow i = j$.

   B. $(\boldsymbol{\sigma'(i) = \sigma'(j) \neq k})$ then $i, j \neq l$ [otherwise we have $\sigma'(i) = \sigma'(l) = k = \sigma'(l) = \sigma'(j)$ by the definition of σ']. So $i, j \in \{1, \ldots, n\} \setminus \{l\}$ and thus $\sigma(i) = \sigma'(i) = \sigma'(j) = \sigma(j) \underset{\sigma \text{ is bijective}}{\Rightarrow} i = j$

   proving injectivity. □

Consider now ${\sigma'}^{[k]} \in P_{n+1}$ and take $j \in \{1, \ldots, n+1\}$ then we have the following cases to consider

i. $(\boldsymbol{j = n+1})$ then

$$\begin{aligned} {\sigma'}^{[k]}(j) &= {\sigma'}^{[k]}(n+1) \\ &= k \\ &= \sigma(n+1) \\ &= \sigma(j) \end{aligned}$$

ii. $(\boldsymbol{j = l})$ then as $\sigma'(l) = k \Rightarrow l = {\sigma'}^{-1}(k)$

$$\begin{aligned} {\sigma'}^{[k]}(j) &= {\sigma'}^{[k]}(l) \\ &= n+1 \\ &= \sigma(l) \\ &= \sigma(j) \end{aligned}$$

iii. $(\boldsymbol{j\in\{1,\ldots,n\}\backslash\{l\}})$ then as $j\neq l=\sigma'^{-1}(k), j\neq n+1$ we have

$$\begin{aligned}\sigma'^{[k]}(j) &= \sigma'(j)\\ &= \sigma(j)\end{aligned}$$

this proves that $\sigma'^{[k]}=\sigma$ and thus $\sigma=\Delta(k,\sigma')$ proving surjectivity. □

We are now finally ready to prove that P_n is a finite set and count its number of elements.

Theorem 10.70. *Let $n\in\mathbb{N}_0$ then P_n is finite and $\#(P_n)=n!$*

Proof. First if $n=0$ then $P_n=P_0$ and $\emptyset:\emptyset\to\emptyset$ is the only bijection in P_n and then $\#(P_n)=1=0!$. Assume now that $n\in\mathbb{N}$ then we prove by induction that P_n is finite with $\#(P_n)=n!$. So let $B=\{n\in\{1,\ldots\}|P_n\,\text{is finite and}\,\#(P_n)=(n)!\}$, we have then:

1. If $n=1$ then $P_n=P_1$ and the only bijection between $S_1=\{0\}$ and $S_1=\{0\}$ is $1_{\{0\}}$ so P_1 is finite and $\#(P_n)=\#(P_1)=1=1!$ so $0\in B$
2. Assume that $n\in B$ then P_n is finite and $\#(P_n)=(n)!$. Then by the previous theorem there exists a bijection $\Delta:\{1,\ldots,n+1\}\times P_n\to P_{n+1}$. As $\{1,\ldots,n+1\}$ is finite with $\#(\{1,\ldots,n+1\})=n+1$ and P_n is finite and $\#(P_n)=n$ we have by 5.44 that $\{1,\ldots,n+1\}\times P_n$ is finite and $\#(\{1,\ldots,n+1\}\times P_n)=(n+1)\cdot(n)!=(n+1)!$. So there exists a bijection $b:(n+1)!\to\{1,\ldots,n+1\}\times P_n$ and thus a bijection $\Delta\circ b:(n+1)!\to P_{n+1}$. So P_{n+1} is finite and $\#(P_{n+1})=(n+1)!$ or $n+1\in B$

By mathematical induction (see 4.77) we have then $B=\{1,\ldots\}=\mathbb{N}$ □

We define now a special permutation called a transposition.

Definition 10.71. *If $n\in\mathbb{N}$ and $i,j\in\{1,\ldots,n\}$ then $(i\leftrightarrow_n j)\in P_n$ is defined by (see 10.16)*

$$\begin{aligned}\big(i\leftrightarrow_n j\big)(k) &= k\,\text{if}\,k\in\{1,\ldots,n\}\backslash\{i,j\}\\ &= j\,\text{if}\,k=i\\ &= i\,\text{if}\,k=j\end{aligned}$$

*is called a transposition. A transposition $(i\leftrightarrow_n j)$ is **strict** if $i\neq j$*

Note 10.72. It is trivially to prove the following statements

1. If $i=j$ then $\big(i\leftrightarrow_n j\big)=1_{\{1,\ldots,n\}}$
2. $(i\leftrightarrow_n j)=(j\leftrightarrow_n i)$
3. $(i\leftrightarrow_n j)\circ(i\leftrightarrow_n j)=1_{\{1,\ldots,n\}}$

Definition 10.73. *As $\langle P_n,\circ\rangle$ forms a group, giving a finite family of $\{\sigma_i\}_{i\in\{1,\ldots,n\}}$ of permutations in P_n we have $\sum_{i=1}^n\sigma_i$ defined ($\circ$ being the sum operator) we use however the notation $(\sigma_1\circ\ldots\circ\sigma_n)$ instead of $\sum_{i=1}^n\sigma_i$*

Note 10.74. As $(\sigma_1 \circ \ldots \circ \sigma_n) = \sum_{i=1}^{n} \sigma_i = \sum_{i=0}^{n-1} \sigma_{i+1}$ we have

$$(\sigma_1 \circ \ldots \circ \sigma_1) \quad = \quad \sum_{i=0}^{0} \sigma_{i+1} = \sigma_1$$

if $n > 1$ then $n = (n-1)+1$ and $(\sigma_1 \circ \ldots \circ \sigma_n)$

$$= \sum_{i=0}^{(n-1)+1} \sigma_{i+1} = \left(\sum_{i=0}^{n-1} \sigma_{i+1}\right) \circ \sigma_{(n-1)+1}$$

$$= (\sigma_1 \circ \ldots \circ \sigma_{n-1}) \circ \sigma_n$$

if $n > 1$ then $(n-1)+1$ and $(\sigma_1 \circ \ldots \circ \sigma_n)$

$$= \sum_{i=\circ}^{(n-1)+1} \sigma_{i+1}$$

$$\underset{10.26, m=0}{=} \left(\sum_{i=0}^{0} \sigma_{i+1}\right) \circ \left(\sum_{i=1}^{(n-1)+1} \sigma_{i+1}\right)$$

$$= \sigma_1 \circ (\sigma_2 \circ \ldots \circ \sigma_n)$$

Summarizing we have that

$$(\sigma_1 \circ \ldots \sigma_1) \;=\; \sigma_1$$

and if $n > 1$ then

$$\begin{aligned}(\sigma_1 \circ \ldots \circ \sigma_n) &= (\sigma_1 \circ \ldots \sigma_{n-1}) \circ \sigma_n \\ &= \sigma_1 \circ (\sigma_2 \circ \ldots \circ \sigma_n)\end{aligned}$$

Lemma 10.75. *If $m, n \in \mathbb{N}$ and $\{\sigma_i\}_{i \in \{1,\ldots,m\}}$ a family of permutations in P_n then we have that $(\sigma_1 \circ \ldots \circ \sigma_m)^{[n+1]} = \left(\sigma_1^{[n+1]} \circ \ldots \circ \sigma_m^{[n+1]}\right)$*

Proof. First we prove that if $\sigma, \tau \in P_n$ that $(\sigma \circ \tau)^{[n+1]} = \sigma^{[n+1]} \circ \tau^{[n+1]}$, so if $i \in \{1, \ldots, n+1\}$ then we have the following cases:

1. $(\boldsymbol{i = n+1})$

$$\begin{aligned}(\sigma \circ \tau)^{[n+1]}(i) &= n+1 \\ &= \sigma^{[n+1]}(n+1) \\ &= \sigma^{[n+1]}(\tau^{[n+1]}(n+1)) \\ &= (\sigma^{[n+1]} \circ \tau^{[n+1]})(n+1)\end{aligned}$$

2. $(\boldsymbol{i \in \{1, \ldots, n\}})$

$$\begin{aligned}(\sigma \circ \tau)^{[n+1]}(i) &= (\sigma \circ \tau)(i) \\ &= \sigma(\tau(i)) \\ &\underset{\tau(i) \in \{1,\ldots,n\}}{=} \sigma^{[n+1]}(\tau(i)) \\ &\underset{i \in \{1,\ldots,n\}}{=} \sigma^{[n+1]}(\tau^{[n+1]}(i)) \\ &= (\sigma^{[n+1]} \circ \tau^{[n+1]})(i)\end{aligned}$$

proving that indeed $(\sigma\circ\tau)^{[n+1]}=\sigma^{[n+1]}\circ\tau^{[n+1]}$

We prove this by mathematical induction (see 4.77) so let $B=\left\{m\in\{1,\ldots\}|(\sigma_1\circ\ldots\circ\sigma_n)^{[n+1]}=\left(\sigma_1^{[n+1]}\circ\ldots\circ\sigma_n^{[n+1]}\right)\right\}$ then we have

1. $(\sigma_1\circ\ldots\sigma_1)^{[n+1]}=\sigma_1^{[n+1]}=\left(\sigma_1^{[n+1]}\circ\ldots\circ\sigma_1^{[n+1]}\right)\Rightarrow 1\in B$
2. if $m\in B$ then we have $(\sigma_1\circ\ldots\sigma_{m+1})^{[n+1]}=((\sigma_1\circ\ldots\circ\sigma_m)\circ\sigma_{m+1})^{[n+1]}=(\sigma_1\circ\ldots\sigma_m)^{[n+1]}\circ\sigma_{m+1}^{[n+1]}\underset{m\in B}{=}\left(\sigma_1^{[n+1]}\circ\ldots\circ\sigma_m^{[n+1]}\right)\circ\sigma_{m+1}^{[n+1]}=\left(\sigma_1^{[n+1]}\circ\ldots\circ\sigma_m^{[n+1]}\right)$

so $B=\{1,\ldots\}$ proving our theorem.

$\square$

Lemma 10.76. *If $n\in\mathbb{N}$ and $i,j\in\{1,\ldots,n\}$ then $(i\leftrightarrow_{n+1}j)=((i\leftrightarrow_n j))^{[n+1]}$. Note that if $(i\leftrightarrow_n j)$ is a strict $(i\leftrightarrow_n j)$ transposition $(i\neq j)$ then trivially $((i\leftrightarrow_n j))^{[n+1]}$ is a strict transposition.*

Proof. If $k\in\{1,\ldots,n+1\}$ then we have the following cases to consider

1. $\boldsymbol{(k=n+1)}$

$$\begin{array}{rcl}(i\leftrightarrow_{n+1}j)(k) & = & (i\leftrightarrow_{n+1}j)(n+1)\\ & \underset{i,j\in\{1,\ldots,n\}\Rightarrow i,j\neq n+\cdots+1}{=} & n+1\\ & = & ((i\leftrightarrow_n j))^{[n+1]}(n+1)\\ & = & ((i\leftrightarrow_n j))^{[n+1]}(k)\end{array}$$

2. $\boldsymbol{(k=i)}$

$$\begin{array}{rcl}(i\leftrightarrow_{n+1}j)(k) & = & j\\ & = & (i\leftrightarrow_n j)(i)\\ & \underset{i\in\{1,\ldots,n\}\Rightarrow i\neq n+1}{=} & ((i\leftrightarrow_n j))^{[n+1]}(i)\\ & = & ((i\leftrightarrow_n j))^{[n+1]}(k)\end{array}$$

3. $\boldsymbol{(k=j)}$

$$\begin{array}{rcl}(i\leftrightarrow_{n+1}j)(k) & = & i\\ & = & (i\leftrightarrow_n j)(j)\\ & \underset{j\in\{1,\ldots,n\}\Rightarrow j\neq n+1}{=} & ((i\leftrightarrow_n j))^{[n+1]}(j)\\ & = & ((i\leftrightarrow_n j))^{[n+1]}(k)\end{array}$$

4. $\boldsymbol{(k\neq i,j,n+1)}$

$$\begin{array}{rcl}(i\leftrightarrow_{n+1}j)(k) & \underset{k\neq i,j}{=} & k\\ & \underset{k\neq i,j,n+1}{=} & (i\leftrightarrow_n j)(k)\\ & \underset{k\neq n+1}{=} & ((i\leftrightarrow_n j))^{[+1]}(k)\end{array}$$

So having considered all the cases we conclude that $\left(i\leftrightarrow_n j\right)=((i<))^{n+1}$ $\square$

By using mathematical induction we can extend the above lemma

Lemma 10.77. *Given $n,m\in\mathbb{N},0<m$ and $\{(k_i\leftrightarrow_n l_i)\}_{i\in\{1,\ldots,m\}}$ is a family of transpositions then $\left(\left(k_1\leftrightarrow_{n+1} l_1\right)\circ\ldots\circ(k_m\leftrightarrow_{n+1} l_m)\right)=(((k_1\leftrightarrow_n l_1)\circ\ldots\circ(k_m\leftrightarrow_n l_m)))^{[n+1]}$*

Proof. $(((k_1\leftrightarrow_n l_1)\circ\ldots\circ(k_m\leftrightarrow_n l_m)))^{[n+1]}\underset{10.75}{=}((k_1\leftrightarrow_n l_1)^{[n+1]}\circ\ldots\circ(k_m\leftrightarrow_n l_m)^{[n+1]})\underset{\text{previous lemma}}{=}((k_1\leftrightarrow l_1)^{[n+1]}\circ\ldots(k_m,l_m)^{[n+1]})$ $\square$

Theorem 10.78. *Given $n\in\mathbb{N}$, $n>1$ then if $\sigma\in P_n$ there exists a $\{(k_i\leftrightarrow_n l_i)\}_{i\in\{1,\ldots,m\}}$, $\forall i\in\{1,\ldots,m\}\vDash k_i\neq l_i$ and $k_i,l_i\in\{1,\ldots,n\}$ (a family of strict transpositions) such that $\sigma=((k_i\leftrightarrow_n l_i)\circ\ldots\circ(k_m\leftrightarrow_n l_m))$*

Proof. We prove this by induction on n so let $B=\{n\in\{2,\ldots\}|\forall\sigma\in P_n$ there exists a $\{(k_i\leftrightarrow_n l_i)\}_{i\in\{1,\ldots,m\}}$ of strict transpositions so that $((k_1\leftrightarrow_n l_1)\circ\ldots\circ(k_m\leftrightarrow_n l_m))\}$. We have then :

1. If $n=2$ then we have the following possibilities for σ

   a. σ is defined by

   $$\begin{aligned}\sigma(1)&=1\\ \sigma(2)&=2\end{aligned}$$

   so $\sigma=i_{\{1,2\}}$ and we can write $\sigma=i_{\{1,2\}}=(1\leftrightarrow_2 2)\circ(1\leftrightarrow_2 2)$ so if $\{(k_i,l_i)\}_{i\in\{1,\ldots 2\}}=\{(1\leftrightarrow_2 2),(1\leftrightarrow_2 2)\}$ then $\sigma=((k_1\leftrightarrow_2 l_1)\ldots(k_2\leftrightarrow_2 l_2))$

   b. σ is defined by

   $$\begin{aligned}\sigma(1)&=2\\ \sigma(2)&=1\end{aligned}$$

   and then clearly $\sigma=(1\leftrightarrow_2 2)$ so if $\{(k_i,l_i)\}_{i\in\{1,\ldots 2\}}=\{(1\leftrightarrow_2 2)\}$ then $\sigma=((k_1\leftrightarrow_2 l_1)\ldots(k_2\leftrightarrow_2 l_2))$

   proving that $2\in B$

2. Assume that $n\in B$ then given $\sigma\in P_{n+1}$ we can consider the following cases

   a. $\boldsymbol{(\sigma(n+1)=n+1)}$ then we have $\sigma=(\sigma_{|\{1,\ldots,n\}})^{[n+1]}$ (see definition of $\sigma^{[i]}$) and as $\sigma_{|\{1,\ldots,n\}}\in P_n$ we have as $n\in B$ that there exists a family of strict transpositions in P_n $\{(k_i\leftrightarrow_n l_i)\}_{i\in\{1,\ldots,m\}}$ so that $\sigma_{|\{1,\ldots,n\}}=((k_1\leftrightarrow_n l_1)\circ\ldots\circ(k_m\leftrightarrow_n l_m))$ and thus $\sigma=(((k_1\leftrightarrow_n l_1)\circ\ldots\circ(k_m\leftrightarrow_n l_m)))^{[n+1]}\underset{\text{previous lemma}}{=}((k_1\leftrightarrow_{n+1} l_1)\circ\ldots\circ(k_m\leftrightarrow_{n+1} l_m))$. So we have found a family of strict transpositions $\{(k_i\leftrightarrow_{n+1} l_i)\}_{i\in\{1,\ldots,m\}}$ in P_{n+1} so that $\sigma=((k_1\leftrightarrow_{n+1} l_1)\circ\ldots\circ(k_m\leftrightarrow_{n+1} l_m))$ and thus $n+1\in B$

b. $(\boldsymbol{\sigma(n+1)\neq n+1})$ Take then $i=\sigma(n+1)\Rightarrow i\neq n+1$ then we have $\sigma=i_{\{1,\ldots,n+1\}}\circ\sigma=((i\leftrightarrow_{n+1}n+1)\circ(i\leftrightarrow_{n+1}n+1))\circ\sigma=(i\leftrightarrow_{n+1}n+1)\circ((i\leftrightarrow_{n+1}n+1)\circ\sigma)=(i\leftrightarrow_{n+1}n+1)\circ\sigma'$ where $\sigma'=\big(i\leftrightarrow_{n+1}n+1\big)\circ\sigma$. We have then that $\sigma'(n+1)=(i\leftrightarrow_{n+1}n+1)(\sigma(n+1))=\big(i\leftrightarrow_{n+1}n+1\big)(i)=n+1$. Just like in (a) we have then that $\sigma'=(\sigma'_{|\{1,\ldots,n\}})^{[n+1]}$ and as $\sigma'_{|\{1,.,n\}}\in P_n$ we have as $n\in B$ that there exists a family of strict transpositions $\{(k_i\leftrightarrow_n l_i)\}_{i\in\{1,\ldots,m\}}$ such that $\sigma'_{|\{1,\ldots,n\}}=((k_1\leftrightarrow_n l_1)\circ\ldots\circ(k_m\leftrightarrow_n l_m))$ and thus $\sigma'=(\sigma'_{|\{1,\ldots,n\}})^{[n+1]}=(((k_1\leftrightarrow_n l_1)\circ\ldots\circ(k_m\leftrightarrow_n l_m)))^{[n+1]}=((k_1\leftrightarrow_{n+1}l_1)\circ\ldots\circ(k_m\leftrightarrow_{n+1}l_m))$. Define now $f=\{(r_i\leftrightarrow_{n+1}s_i)\}_{i\in\{1,\ldots,m+1\}}$ by $\forall o\in\{1,\ldots,m+1\}$

$$\begin{aligned} f(o) &= (i\leftrightarrow_{n+1}n+1)\text{ if }o=1\ (a\text{ strict transposition})\\ &= (k_{o-1}\leftrightarrow_{n+1}l_{o-1})\text{ if }o\in\{2,\ldots,m+1\}\end{aligned}$$

then $((r_1\leftrightarrow_{n+1}s_1)\circ\ldots\circ(r_{m+1}\leftrightarrow_{n+1}s_{m+1}))=(r_1\leftrightarrow_{n+1}s_1)\circ((r_2\leftrightarrow_{n+1}s_2)\circ\ldots\circ(r_{m+1}\leftrightarrow_{n+1}s_{m+1}))=(i\leftrightarrow_{n+1}n+1)\circ((k_1\leftrightarrow_{n+1}l_1)\circ\ldots\circ(k_m\leftrightarrow_{n+1}l_m))=(i\leftrightarrow_{n+1}n+1)\circ\sigma'=\sigma$ proving that $n+1\in B$

Using mathematical induction (see 4.77) we have $B=\{2,\ldots\}$ and thus if $n\in\mathbb{N},n>1$ we have $n\in B$ and thus if $\sigma\in P_n$ then we can write it as a combination of strict transpositions. □

Theorem 10.79. *Given $n\in\mathbb{N}$, $\{x_i\}_{i\in\{1,\ldots n\}}$ a family of elements in X and $\emptyset\neq Y\subseteq\{x_i|i\in\{1,\ldots,n\}\}$ then there exists a $m\in\{1,\ldots,n\}$ and a permutation $\sigma\in P_n$ such that $\forall i\in\{1,\ldots,m\}$ we have $x_{\sigma_i}\in Y$ and $\forall i\in\{m+1,\ldots,n\}$ we have $x_{\sigma_i}\notin Y$ [and thus $x_{\sigma_i}\in X\setminus Y$]*

Proof. We prove this by induction so let $A=\{n\in\{1,\ldots\}|\text{if }\{x_i\}_{i\in\{1,\ldots,n\}}$ is a family of elements of a set X and $\emptyset\neq Y\subseteq\{x_i|i\in\{1,\ldots,n\}\}$ then there exists a $m\in\{1,\ldots,n\}$ and a permutation $\sigma\in P_n$ such that $\forall i\in\{1,\ldots,m\}$ we have $x_{\sigma_i}\in Y$ and $\forall i\in\{m+1,\ldots,n\}$ we have $x_{\sigma_i}\notin Y\}$ we have then :

1. If $n=1$ then $\{x_i\}_{i\in\{1,\ldots,1\}}$ is a function from $\{1\}\rightarrow X$ so $\{x_i|i\in\{1,\ldots,1\}\}=\{x_1\}$ and if $\emptyset\neq Y\subseteq\{x_1\}$ we have $Y=\{x_1\}$ and we can that $m=1$ and $\sigma=1_{\{1\}}$ so that if $i\in\{1\ldots1\}\Rightarrow i=1\Rightarrow x_i=x_1\in\{x_1\}=Y$ and if $i\in\{2,\ldots,1\}=\emptyset$ we have that $x_{\sigma_i}\notin Y$ is full filled vacuously. So $1\in A$

2. Assume that $n\in A$ then if $\{x_i\}_{i\in\{1,\ldots,n+1\}}$ is a family of elements of X and $\emptyset\neq Y\subseteq\{x_i|i\in\{1,\ldots,n+1\}\}$. There are two cases to consider:

   a. $(\boldsymbol{Y=\{x_i|i\in\{1,\ldots,n+1\}\}})$ In this case we can take $m=n+1$ and $\sigma=1_{\{1,\ldots,n+1\}}$ and then $\forall i\in\{1,\ldots,m\}=\{1,\ldots,n+1\}$ we have $x_{\sigma_i}=x_i\in\{x_i|i\in\{1,\ldots,n+1\}\}=Y$ and if $i\in\{m+1,\ldots,n+1\}=\emptyset$ then $x_{\sigma_i}\notin Y$ is full filled vacuously. So we conclude that $n+1\in A$

b. $\boldsymbol{(Y \subset \{x_i | i \in \{1, \ldots, n+1\}\})}$ then there exists a $y \in X$ such that $y \notin Y$. As $y \in X$ there exists a $j \in \{1, \ldots, n+1\}$ such that $x_j = y$. Take then $\sigma' \in P_{n+1}$ to be $\sigma' = (j \leftrightarrow_{n+1} n+1)$. We have then for $\{x_{\sigma'_i}\}_{i \in \{1, \ldots, n+1\}}$ that $x_{\sigma'_{n+1}} = x_{(j \leftrightarrow_{n+1} n+1)(n+1)} = x_j = y \notin Y$. This means that for $\{x_{\sigma'_i}\}_{i \in \{1, \ldots, n\}}$ the restriction (as a function) of $\{x_{\sigma'_i}\}_{i \in \{1, \ldots, n+1\}}$ to $\{1, \ldots, n\}$ we have that if $z \in Y$ then from $Y \subseteq \{x_i | i \in \{1, \ldots, n+1\}\}$ that there exists a $i \in \{1, \ldots, n+1\}$ such that $z = x_i$, as σ' is a bijection there exists a $k \in \{1, \ldots, n+1\}$ such that $i = \sigma'(k)$, now if $k = n+1$ then $i = \sigma'(k) = \sigma'(n+1) = j \Rightarrow z = x_i = x_j = y \notin Y$ a contradiction, so $k \in \{1, \ldots, n+1\}$ and thus $z \in \{x_{\sigma'_i} | i \in \{1, \ldots, n\}\}$ proving that $\emptyset \neq Y \subseteq \{x_{\sigma'_i} | i \in \{1, \ldots, n\}\}$. Using the fact that $n \in A$ there exists a $m \in \{1, \ldots, n\}$ and a $\sigma'' \in P_n$ [$\sigma'' : \{1, \ldots, \} \to \{1, \ldots, n\}$ is a bijection] such that $\forall i \in \{1, \ldots m\}$ we have $x_{\sigma'_{\sigma''_i}} \in Y$ and $\forall i \in \{m+1, \ldots n\}$ we have $x_{\sigma'_{\sigma''_i}} \notin Y$. Using the bijection $i_{n+1} : \{n+1\} \to \{n+1\}$ we construct (see 2.43) the bijection $\sigma''' = \sigma'' \bigcup i_{\{n+1\}} : \{1, \ldots, n+1\} \to \{1, \ldots, n+1\}$ so $\sigma'' \in P_{n+1}$. Take then $\sigma = \sigma' \circ \sigma'''$ then $\sigma \in P_{n+1}$ and if $i \in \{1, \ldots, m\}$ we have $\sigma(i) = \sigma'(\sigma'''(i)) \underset{\{1, \ldots, m\} \subseteq \{1, \ldots, n\}}{=} \sigma'(\sigma''(i))$ and thus $x_{\sigma_i} = x_{\sigma'_{\sigma''_i}} \in Y$. Also if $i \in \{m+1, \ldots, n+1\}$ then we have either

i. $\boldsymbol{(i = n+1)}$ then $\sigma(i) = \sigma'(\sigma'''(n+1)) = \sigma'(n+1) = j \Rightarrow x_{\sigma_i} = y \notin Y$

ii. $\boldsymbol{(i \neq n+1)}$ then $i \in \{m+1, \ldots, n\}$ and $\sigma(i) = \sigma'(\sigma'''(i)) \underset{\{m+1, \ldots, n\} \subseteq \{1, \ldots, n\}}{=} \sigma'(\sigma''(i))$ and thus $x_{\sigma_i} = x_{\sigma'_{\sigma''_i}} \notin Y$

Having found our $m \in \{1, \ldots, n\} \subseteq \{1, \ldots, n+1\}$ and $\sigma \in P_{n+1}$ we can say that $n+1 \in A$

Using mathematical induction (see 4.77) we have then that $A = \{1, \ldots.\} = \mathbb{N}$ proving the theorem. □

Definition 10.80. *If X is a set, $n \in \mathbb{N}$ define then the family $\{X_i\}_{i \in \{1, \ldots, n\}}$by $\forall i \in \{1, \ldots, n\}$ we have $X_i = X$ then $X^n = \prod_{i \in \{1, \ldots, n\}} X_i$. Using the definition of a 2.77 we have then $X^n = \{f | f : \{1, \ldots, n\} \to X \text{ is a function}\}$. If $x \in X^n$ then $\forall i \in \{1, \ldots, n\}$ we note $x(i)$ as x_i this is the reason why another notation for the elements of X^n is $(x_1, \ldots, x_n)$ (especially if we have a formula to calculate $x(x) = x_i$. If $x \in X^n$and $\sigma \in P_n$ then $x \circ \sigma : \{1, \ldots, n\} \to X$ and thus $x \circ \sigma \in X^n$ we note this element as $(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$ or $(x_{\sigma_1}, \ldots, x_{\sigma_n})$*

Example 10.81. If $n = 1$ then $X^1 = \{f : f : \{1\} \to x\}$ we can now find a trivial bijection between X and X^1 given $x \in X$ from $\{1\}(x) : \{1\} \to X$ by $\{1\}(x)(1) = x$

1. **(injectivity)** If $\{1\}(x) = \{1\}(y) \Rightarrow x = \{1\}(x)(1) = \{1\}(y) = y \Rightarrow x = y$
2. **(surjectivity)** If $f \in X^1$ take then $x = f(1)$ then $\{1\}(x)(1) = x = f(1) \Rightarrow \{1\}(x) = f$

Definition 10.82. *Let X, Y be sets, $n \in \mathbb{N}$, $f: X^n \rightarrow Y$ a function and $\sigma \in P_n$ then $\sigma f: X^n \rightarrow Y$ is defined by $\forall x \in X^n$ we have $(\sigma f)(x) = f(x \circ \sigma)$ or in other notation $(\sigma f)(x_1, ..., x_n) = f(x_{\sigma_1}, ..., x_{\sigma_n})$*

Theorem 10.83. *Let X, Y be sets, $\sigma, \tau \in P_n$ and $f: X^n \rightarrow Y$ then $\tau(\sigma f) = (\tau \circ \sigma) f$*

Proof. If $x \in X^n$ then

$$\begin{array}{rcl}(\tau(\sigma f))(x) & = & (\sigma f)(x \circ \tau) \\ & = & f((x \circ \tau) \circ \sigma) \\ & = & f(x \circ (\tau \circ \sigma)) \\ & = & ((\tau \circ \sigma) f)(x)\end{array}$$

and thus we conclude that $\tau(\sigma f) = (\tau \circ \sigma) f$ □

Definition 10.84. *Given $n \in \mathbb{N}$ then $\Phi_n: \mathbb{Z}^n \rightarrow \mathbb{Z}$ is defined by $\Phi_n(x) = \prod_{(i,j) \in \{(i,j) \in \{1,...,n\} \times \{1,...,n\} | i < j\}} (x_i - x_j)$*

Example 10.85.

1. $\Phi_1(2) = 1$ as $\{(i, j) \in \{1\} \times \{1\} | i < j\} = \emptyset$ (here 1 is neutral element in $\langle \mathbb{Z}, \cdot \rangle$
2. $\Phi_2(1, 2) = \prod_{(i,j) \in \{(1,2)\}} (x_i - x_j) = (1-2)$
3. $\Phi_3(1, 2, 3) = (1-2) \cdot (1-3) \cdot (2-3) = 2$

Theorem 10.86. *Given $n \in \mathbb{N}, 1 < n$ then if $k, l \in \{1, ..., n\}, k \neq l$ and $\sigma = (k \leftrightarrow_n l)$ is a strict transposition then $\sigma \Phi_n = (-1) \cdot \Phi_n$*

Proof. First define $I = \{(i, j) \in \{1, ..., n\} \times \{1, ..., n\} | i < j\}$ which is finite, then if $k \neq l$ we have the following possibilities:

1. $(\boldsymbol{k < l})$ If $(i, j) \in I \Rightarrow i < j$ then we have the following excluding possibilities
   a. $(\boldsymbol{i = k, j = l})$ $\Rightarrow (i, j) \in \{(k, l)\} = I_1 \subseteq I$
   b. $(\boldsymbol{i = k, j \neq l})$ then we have the following sub-cases
      i. $(\boldsymbol{j < l})$ $\underset{k=i<j}{\Rightarrow} k < j < l \Rightarrow (i, j) \in \{(k, j) \in I | k < j < l\} = I_2 \subseteq I$
      ii. $(\boldsymbol{l < j})$ $\Rightarrow (i, j) \in \{(k, j) \in I | l < j\} = I_3 \subseteq I$
   c. $(\boldsymbol{i \neq k})$ with the following sub-cases
      i. $(\boldsymbol{i = l})$ $\underset{i<j}{\Rightarrow} (i, j) \in \{(l, j) \in I | l < j\} = I_4 \subseteq I$
      ii. $(\boldsymbol{i \neq l})$ with the following sub-cases
         A. $(\boldsymbol{j = k})$ $\underset{i<j}{\Rightarrow} i < k \Rightarrow (i, j) \in \{(i, k) \in I | i < k\} = I_5 \subseteq I$
         B. $(\boldsymbol{j = l})$ then we have the following sub-cases
            1. $(\boldsymbol{i < k})$ $\Rightarrow (i, j) \in \{(i, l) \in I | i < k\} = I_6 \subseteq I$

2. $\boldsymbol{(k < i)}$ $\underset{i<j=l}{\Rightarrow}(i,j) \in \{(i,l) \in I | k < i < l\} = I_7 \subseteq I$

C. $\boldsymbol{(j \neq l, k)}$ $\Rightarrow (i,j) \in \{(i,j) \in I | i \neq k, l \wedge j \neq k, l\} = I_8 \subseteq I$

proving that $I \subseteq \bigcup_{k \in \{1,\ldots,8\}} I_k \subseteq I \Rightarrow I = \bigcup_{k \in \{1,\ldots,8\}} I_k$. Now for the intersections of the form $I_i \bigcap I_j, i \neq j$ we have the following cases to consider:

a. $\boldsymbol{((i,j) \in I_1 \bigcap I_2)}$ then $i = k \wedge j = l \wedge j < l \Rightarrow$ a contradiction

b. $\boldsymbol{((i,j) \in I_1 \bigcap I_3)}$ then $i = k \wedge j = l \wedge l < j \Rightarrow$ a contradiction

c. $\boldsymbol{((i,j) \in I_1 \bigcap I_4)}$ then $i = k \wedge j = l \wedge l < j \Rightarrow$ a contradiction

d. $\boldsymbol{((i,j) \in I_1 \bigcap I_5)}$ then $i = k \wedge j = l \wedge i < k \Rightarrow$ a contradiction

e. $\boldsymbol{((i,j) \in I_1 \bigcap I_6)}$ then $i = k \wedge j = l \wedge i < k \Rightarrow$ a contradiction

f. $\boldsymbol{((i,j) \in I_1 \bigcap I_7)}$ then $i = k \wedge j = l \wedge k < i < l \Rightarrow$ a contradiction

g. $\boldsymbol{((i,j) \in I_1 \bigcap I_8)}$ then $i = k \wedge j = l \wedge i \neq k \Rightarrow a$ contradiction

h. $\boldsymbol{((i,j) \in I_2 \bigcap I_3)}$ then $i = k \wedge k < j < l \wedge l < j \Rightarrow a$ contradiction

i. $\boldsymbol{((i,j) \in I_2 \bigcap I_4)}$ then $i = k \wedge k < j < l \wedge i = l \wedge l < j \Rightarrow$ a contradiction

j. $\boldsymbol{((i,j) \in I_2 \bigcap I_5)}$ then $i = k \wedge k < j < l \wedge k = j \Rightarrow$ a contradiction

k. $\boldsymbol{((i,j) \in I_2 \bigcap I_6)}$ then $i = k \wedge k < j < l \wedge j = l \wedge i < k \Rightarrow$ a contradiction

l. $\boldsymbol{((i,j) \in I_2 \bigcap I_7)}$ then $i = k \wedge k < j < l \wedge j = l \wedge k < i < l \Rightarrow$ a contradiction

m. $\boldsymbol{((i,j) \in I_2 \bigcap I_8)}$ then $i = k \wedge k < j < l \wedge i, j \neq k, l \Rightarrow$ a contradiction

n. $\boldsymbol{((i,j) \in I_3 \bigcap I_4)}$ then $i = k \wedge l < j \wedge i = l \wedge l < j \Rightarrow$ a contradiction

o. $\boldsymbol{((i,j) \in I_3 \bigcap I_5)}$ then $i = k \wedge l < j \wedge i < k \wedge j = k \Rightarrow$ a contradiction

p. $\boldsymbol{((i,j) \in I_3 \bigcap I_6)}$ then $i = k \wedge l < j \wedge j = l \wedge i < k \Rightarrow$ a contradiction

q. $\boldsymbol{((i,j) \in I_3 \bigcap I_7)}$ then $i = k \wedge l < j \wedge j = l \wedge k < i < l \Rightarrow$ a contradiction

r. $\boldsymbol{((i,j) \in I_3 \bigcap I_8)}$ then $i = k \wedge l < j \wedge i, j \neq k, l \Rightarrow$ a contradiction

s. $\boldsymbol{((i,j) \in I_4 \bigcap I_5)}$ then $i = l \wedge l < j \wedge j = k \wedge i < k \Rightarrow l < k$ contradicting $k < l$

t. $\boldsymbol{((i,j) \in I_4 \bigcap I_6)}$ then $i = l \wedge l < j \wedge j = l \wedge i < k \Rightarrow l < l$ a contradiction

u. $\boldsymbol{((i,j) \in I_4 \bigcap I_7)}$ then $i = l \wedge l < j \wedge j = l \wedge k < i < l \Rightarrow l < l$ a contradiction

v. $\boldsymbol{((i,j) \in I_4 \bigcap I_8)}$ then $i = l \wedge l < j \wedge i, j \neq k, l \Rightarrow$ s contradiction

w. $\boldsymbol{((i,j) \in I_5 \bigcap I_6)}$ then $j = k \wedge i < k \wedge j = l \wedge i < k \Rightarrow k = l$ contradicting $k < l$

x. $\boldsymbol{((i,j) \in I_5 \bigcap I_7)}$ then $j = k \wedge i < k \wedge j = l \wedge k < i < l \Rightarrow i < i$ a contradiction

y. $\boldsymbol{((i,j) \in I_5 \bigcap I_8)}$ then $j = k \wedge i < k \wedge i, j \neq k, l \Rightarrow$ a contradiction

z. $\boldsymbol{((i,j)\in I_6\bigcap I_7)}$ then $j=l\wedge i<k\wedge j=l\wedge k<i<l\Rightarrow i<i$ a contradiction

aa. $\boldsymbol{((i,j)\in I_6\bigcap I_8)}$ then $j=l\wedge i<k\wedge i,j\neq k,l\Rightarrow$ a contradiction

ab. $\boldsymbol{((i,j)\in I_7\bigcap I_8)}$ then $j=l\wedge k<i<l\wedge i,j\neq k,l\Rightarrow$ a contradiction

So we have (taking commutativity of $\bigcap$ in account) that $\forall i,j\in\{1,...,n\}$ with $i\neq j$ we have $I_i\bigcap I_j=\emptyset$. Using 10.46 we have that

$$\begin{aligned}\Phi(x_1,...,x_n) &= \left(\prod_{k=1}^{8}\left(\prod_{(i,j)\in I_k}(x_i-x_j)\right)\right)\\ &= Q_1\cdot Q_2\cdot Q_3\cdot Q_4\cdot Q_5\cdot Q_6\cdot Q_7\\ \text{here } Q_m &= \prod_{(i,j)\in I_m}(x_i-x_j)\end{aligned}$$

and

$$\begin{aligned}\sigma\Phi(x_1,...,x_n) &= \Phi(x_{\sigma_1},...,x_{\sigma_n})\\ &= \prod_{k=1}^{8}\left(\prod_{(i,j)\in I_k}(x_{\sigma_i}-x_{\sigma_j})\right)\\ &= P_1\cdot P_2\cdot P_3\cdot P_4\cdot P\\ \text{here } P_m &= \prod_{(i,j)\in I_m}(x_{\sigma_i}-x_{\sigma_j})\end{aligned}$$

Now we have the following 8 products to calculate

a. $P_1=\prod_{(i,j)\in I_1}(x_{\sigma_i}-x_{\sigma_j})=(x_{\sigma_k}-x_{\sigma_l})=(x_l-x_k)=-(x_k-x_l)=-\prod_{(i,j)\in I_1}(x_i-x_j)=-Q_1$

b. $P_2=\prod_{(i,j)\in I_2}(x_{\sigma_i}-x_{\sigma_j})=\prod_{(i,j)\in\{(k,j)\in I|k<j<l\}}(x_{\sigma_i}-x_{\sigma_j})=\prod_{(i,j)\in\{(k,j)\in I|k<j<l\}}(x_{\sigma_k}-x_{\sigma j})=\prod_{(i,j)\in\{(k,j)\in I|k<j<l|\}}(x_l-x_j)=\prod_{(i,j)\in I_2}(x_l-x_j)=\prod_{(i,j)\in I_2}A_{(i,j)}$ where $A_{(i,j)}=x_l-x_j$. Now we construct the following function b: $I_7=\{(i,l)\in I|k<i<l\}\rightarrow I_2=\{(k,j)\in I|k<j<l\}$ by $(i,j)\rightarrow b(i,j)=(k,i)$ this is a bijection as

i. **(injectivity)** If $b(i,j)=b(i',j')\underset{(i,j)\in I_7}{\Rightarrow}j=l=j'\wedge(k,i)=(k,i')\Rightarrow i=i'$ and thus $(i,j)=(i',j')$.

ii. **(surjectivity)** If $(i,j)\in I_2\Rightarrow k=i\wedge k<j<l\Rightarrow(i,j)=(k,j)$ and $(j,l)\in I_7$ and $b(j,l)=(k,j)=(i,j)$

Using 10.44 (applied to products) we get $\prod_{(i,j)\in I_2}A_{(i,j)}=\prod_{(i,j)\in I_7}A_{b(i,j)}=\prod_{(i,j)\in I_7}A_{(k,i)}=\prod_{(i,j)\in I_7}(x_l-x_i)=\prod_{(i,j)\in I_7}(-1)\cdot(x_i-x_l)\underset{10.55}{=}\left(\prod_{(i,j)\in I_7}(x_i-x_l)\right)(-1)^{\#(I_7)}\underset{(i,j)\in I_7\Rightarrow j=l}{=}\left(\prod_{(i,j)\in I_7}(x_i-x_j)\right)\cdot(-1)^{\#(I_7)}=Q_7\cdot(-1)^{\#(I_7)}$ so we have proved that

$$P_2=(-1)^{\#(i_7)}\cdot Q_7$$

c. $P_3 = \prod_{(i,j)\in I_3} (x_{\sigma_i} - x_{\sigma_j}) \underset{(i,j)\in I_3 \Rightarrow i=k \text{ and } j\neq l}{=} \prod_{(i,j)\in I_3} (x_l - x_j) = \prod_{(i,j)\in I_3} A_{(i,j)}$ where here $A_{i,j} = x_l - x_j$. Define now $b: I_4 = \{(l,j) \in I | l < j\} \to I_3 = \{(k,j) \in I | l < j\}$ defined by $(i,j) \to (k,j)$ then b is bijective:

   i. **(injectivity)** If $b(i,j) = b(i',j') \underset{(i,j),(i',j')\in I_4}{\Rightarrow} i = l = i' \wedge (k, j) = (k, j') \Rightarrow i = i' \wedge j = j' \Rightarrow (i,j) = (i',j')$

   ii. **(surjectivity)** If $(i,j) \in I_3 \Rightarrow i = k \wedge l < j \Rightarrow (i,j) = (k,j) \wedge (l, j) \in I_4$ and $b(l,j) = (k,j) = (i,j)$

   Using 10.44 we have $\prod_{(i,j)\in I_3} A_{(i,j)} = \prod_{(i,j)\in I_4} A_{b(i,j)} = \prod_{(i,j)\in I_4} A_{(k,j)} = \prod_{(i,j)\in I_4} (x_l - x_j) \underset{(i,j)\in I_4 \Rightarrow i=l}{=} \prod_{(i,j)\in I_4} (x_i - x_j) = Q_4$ so we have proved that

$$P_3 = Q_4$$

d. $P_4 = \prod_{(i,j)\in I_4} (x_{\sigma_i} - x_{\sigma_j}) \underset{(i,j)\in I_4 \Rightarrow i=l \wedge k<l<j}{=} \prod_{(i,j)\in I_4} (x_k - x_j) = \prod_{(i,j)\in I_4} A_{(i,j)}$ where $A_{(i,j)} = x_k - x_j$. Define now $b: I_3 = \{(k,j) \in I | l < j\} \to I_4 = \{(l,j) \in I | l < j\}$ by $(i,j) \to b(i,j) = (l,j)$ then b is bijective:

   i. **(injectivity)** If $b(i,j) = b(i',j') \underset{(i,j),(i',j')\in I_3}{\Rightarrow} i = k = i' \wedge (l, j) = (l, j') \Rightarrow (i,j) = (i',j')$

   ii. **(surjectivity)** If $(i,j) \in I_4 \Rightarrow i = l \wedge l < j \Rightarrow (i,j) = (l,j) \wedge (k, j) \in I_3$ and $b(k,j) = (l,j) = (i,j)$

   Using 10.44 we have then that $\prod_{(i,j)\in I_4} A_{(i,j)} = \prod_{(i,j)\in I_3} A_{b(i,j)} = \prod_{(i,j)\in I_3} A_{(l,j)} = \prod_{(i,j)\in I_3} (x_k - x_j) \underset{(i,j)\in I_3 \Rightarrow i=k}{=} \prod_{(i,j)\in I_3} (x_i - x_j) = Q_3$, so we have proved that

$$P_4 = Q_3$$

e. $P_5 = \prod_{(i,j)\in I_5} (x_{\sigma_i} - x_{\sigma_j}) \underset{(i,j)\in I_5 \Rightarrow j=k \wedge i<k<l}{=} \prod_{(i,j)\in I_5} (x_i - x_l) = \prod_{(i,j)\in I_5} A_{(i,j)}$ where $A_{(i,j)} = x_i - x_l$. Define now $b: I_6 = \{(i,l) \in I | i < k\} \to I_5 = \{(i,k) \in I | i < k\}$ by $(i,j) = b(i,j) = (i,k)$ then b is a bijection :

   i. **(injectivity)** If $b(i,j) = b(i',j') \underset{(i,j),(i',j')\in I_6}{\Rightarrow} j = l = j' \wedge (i, k) = (i',k) \Rightarrow (i,j) = (i',j')$

   ii. **(surjectivity)** If $(i,j) \in I_5 \Rightarrow (i,j) = (i,k) \wedge i < k \Rightarrow (i,l) \in I_6$ and $b(i,l) = (i,k) = (i,j)$

   Using 10.44 we have $\prod_{(i,j)\in I_5} A_{(i,j)} = \prod_{(i,j)\in I_6} A_{b(i,j)} = \prod_{(i,j)\in I_6} A_{(i,k)} = \prod_{(i,j)\in I_6} (x_i - x_l) \underset{(i,j)\in I_6 \Rightarrow j=l}{=} \prod_{(i,j)\in I_6} (x_i - x_j) = Q_6$, so we have proved that

$$P_5 = Q_6$$

f. $P_6 = \prod_{(i,j)\in I_6} (x_{\sigma_i} - x_{\sigma_j}) \underset{(i,j)\in I_6 \Rightarrow j=l \wedge i<k<l}{=} \prod_{(i,j)\in I_6} (x_i - x_k) = \prod_{(i,j)\in I_6} A_{(i,j)}$ where $A_{(i,j)} = x_i - x_k$. Define now $b: I_5 = \{(i,k)\in I | i < k\} \rightarrow I_6 = \{(i,l) | i<k\}$ by $(i,j)\rightarrow b(i,j) = (i,l)$ then b is a bijection :

   i. **(injectivity)** If $b(i,j) = b(i',j') \underset{(i,j),(i',j')\in I_5}{\Rightarrow} j = k = j' \wedge (i, l) = (i', l) \Rightarrow (i,j) = (i',j')$

   ii. **(surjectivity)** If $(i,j)\in I_6 \Rightarrow (i,j) = (i,l) \cap i<k \Rightarrow (i,k)\in I_5$ and $b(i,k) = (i,l) = (i,j)$

   Using 10.44 we have $\prod_{(i,j)\in I_6} A_{(i,j)} = \prod_{(i,j)\in I_5} A_{b(i,j)} = \prod_{(i,j)\in I_5} A_{(i,l)} = \prod_{(i,j)\in I_5} (x_i - x_k) \underset{(i,j)\in I_5 \Rightarrow j=k}{=} \prod_{(i,j)\in I_5} (x_i - x_j) = Q_5$, so we have proved that

$$P_6 = Q_5$$

g. $P_7 = \prod_{(i,j)\in I_7} (x_{\sigma_i} - x_{\sigma_j}) \underset{(i,j)\in I_7 \Rightarrow j=l \wedge k<i<l}{=} \prod_{(i,j)\in I_7} (x_i - x_k) = \prod_{(i,j)\in I_7} A_{(i,j)}$ where $A_{(i,j)} = x_i - x_k$. Define now $b: I_2 = \{(k,j)\in I | k<j<l|\} \rightarrow I_7 = \{(i,l)\in I | k<i<l\}$ by $(i,j)\rightarrow (j,l)$ then b is a bijection:

   i. **(injectivity)** If $b(i,j) = b(i',j') \underset{(i,j),(i',j)\in I_7}{\Rightarrow} i = k = i' \wedge (j, l) = (j', l) \Rightarrow (i,j) = (i',j')$

   ii. **(surjectivity)** If $(i,j)\in I_7 \Rightarrow (i,j) = (i,l) \wedge k<i<l \Rightarrow (k, i)\in I_2$ and $b(k,i) = (i,l) = (i,j)$

   Using 10.44 we have that $\prod_{(i,j)\in I_7} A_{(i,j)} = \prod_{(i,j)\in I_2} A_{b(i,j)} = \prod_{(i,j)\in I_2} A_{(j,l)} = \prod_{(i,j)\in I_2} (x_j - x_k) \underset{(i,j)\in I_2 \Rightarrow i=k}{=} \prod_{(i,j)\in I_2} (x_j - x_i) = \prod_{(i,j)\in I_2} ((-1)\cdot(x_i - x_j)) = \left(\prod_{(i,j)\in I_2} (x_i - x_j)\right)\cdot (-1)^{\#(I_2)} \underset{I_2\approx I_7 \Rightarrow \#(I_2)=\#(I_7)}{=} Q_2\cdot(-1)^{\#(I_7)}$, this proves that

$$P_7 = (-1)^{\#(I_7)}\cdot Q_2$$

h. $P_8 = \prod_{(i,j)\in I_8} (x_{\sigma_i} - x_{\sigma_j}) \underset{(i,j)\in I_8 \Rightarrow i,j\neq k,l}{=} \prod_{(i,j)\in I_8} (x_i - x_j) = Q_8$

So using the above easy but elaborate calculations we have that

$$\begin{aligned}
\sigma\Phi(x_1,\ldots,x_n) &= P_1\cdot P_2\cdot P_3\cdot P_4\cdot P_5\cdot P_6\cdot P_7\cdot P_8\\
&= (-1)\cdot Q_1\cdot(-1)^{\#(I_7)}\cdot Q_7\cdot Q_4\cdot Q_3\cdot Q_6\cdot Q_5\cdot(-1)^{\#(I_7)}\cdot Q_2\cdot\\
&\quad Q_8\\
&= (-1)\cdot Q_1\cdot Q_2\cdot Q_3\cdot Q_4\cdot Q_5\cdot Q_6\cdot Q_7\cdot Q_8\\
&= -\Phi(x_1,\ldots,x_n)
\end{aligned}$$

where we used associativity, commutativity and 6.19.

2. $\boldsymbol{(l<k)}$ Take then $l' = k$ and $k' = l$, then we have $k' < l'$ and if $(k\leftrightarrow_n l) = (l'\leftrightarrow_n k') = (k'\leftrightarrow_n l')$ then we have

$$\begin{aligned}
(k\leftrightarrow_n l)\Phi(x_1,\ldots,x_n) &= (k'\leftrightarrow_n l')\Phi(x_1,\ldots,x_n)\\
&\underset{(\text{use}(1))}{=} -\Phi(x_1,\ldots,x_n)
\end{aligned}$$

□

Lemma 10.87. *If $\{x_i\}_{i\in I}$ is a finite family of elements in $\mathbb{Z}$ where I is a finite set and $\forall i\in I$ we have $x_i\neq 0$ then $\prod_{i\in I} x_i\neq 0$*

Proof. First we prove by induction on $n\in\mathbb{N}$ that if $\{x_i\}_{i\in\{0,\ldots,n-1\}}$ is such that $x_i\neq 0$ then $\prod_{i=0}^{n-1} x_i\neq 0$. So let $X=\{n\in\{1,\ldots\}|\text{If }\{x_i\}_{i\in\{0,\ldots,n-1\}}\text{ is such that }x_i\neq 0\Rightarrow\prod_{i=0}^{n-1}\}$ then

1. $\prod_{i=0}^{1-1} x_i=x_0\neq 0\Rightarrow 1\in X$
2. If $n\in X\Rightarrow\prod_{i=0}^{(n+1)-1}=\prod_{i=0}^{(n-1)+1} x_i=(\prod_{i=0}^{n-1} x_i)\cdot x_n\neq 0$ because $n\in X$ and $\langle Z,+,\cdot\rangle$ is a integral domain. So $n+1\in X$

Using mathematical induction (see 4.77) we have $X=\{1,\ldots\}=\mathbb{N}$.

Now if I is finite and $n=\#(I)$ and $h: S_n\to I$ a bijection then by 10.51 we have $\prod_{i\in I} x_i=\prod_{i=0}^{n-1} x_{h_i}\neq 0$ [previous proof and the fact that $x_{h_i}\neq 0$] □

Theorem 10.88. *If $n\in\mathbb{N}$ then $\forall\sigma\in P_n$ there exists a unique $\varepsilon_\sigma\in\{-1,1\}$ such that $\sigma\Phi=\varepsilon_\sigma\Phi$ (where Φ is as defined in 10.84). Further if $\sigma=((i_1\leftrightarrow j_1)\circ\ldots\circ(i_m\leftrightarrow j_m))$ then $\varepsilon_\sigma=(-1)^m$ (which by 6.19 means that $\varepsilon_\sigma=-1$ if σ is a odd number of strict transpositions and 1 if σ is a even number of strict transpositions.*

Proof.

We do this proof in two stages

1. **(Existence)** If $n=1$ then $\sigma=i_{\{1\}}$ and $\sigma\Phi(x)=\Phi(x\circ\sigma)=\Phi(x)\Rightarrow\sigma\Phi=\Phi$. If $\sigma\in P_n, n\in\{2,\ldots\}$ then by 10.78 $\sigma=((i_1\leftrightarrow_n j_1)\circ\ldots\circ(i_m\leftrightarrow_n j_m))$. We prove now by induction on n that $\sigma\Phi=(-1)^m\Phi$. So let $X=\{m\in\{1,\ldots\}|\text{If }\sigma=((i_1\leftrightarrow_n j_1)\circ\ldots\circ(i_m\leftrightarrow_n j_m))\text{ then }\sigma\Phi=(-1)^m\Phi\}$ then we have
   a. If $n=1\Rightarrow\sigma=(i_1\leftrightarrow_n j_1)$ and using 10.86 we have $\sigma\Phi=-\Phi=(-1)^1\Phi$ so $1\in X$
   b. If $m\in X$ then if $\sigma=((i_1\leftrightarrow_n i_1)\circ\ldots.\circ(i_{m+1}\leftrightarrow_n j_{m+1}))=\sigma'\circ(i_{m+1}\leftrightarrow_n j_{m+1})$ where $\sigma'=((i_1\leftrightarrow_n j_1)\circ\ldots\circ(i_m\leftrightarrow j_m))$ and using 10.83 we have $(\sigma'\circ(i_{m+1}\leftrightarrow_m j_{n+1}))\Phi=\sigma'((i_{m+1}\leftrightarrow_n j_{m+1})\Phi)=\sigma'(-\Phi)=-\sigma'\Phi\underset{n\in X}{=}-(-1)^m\Phi=(-1)^{m+1}\Phi$ proving that $m+1\in X$

   Using induction (see 4.10) we have $X=\{1,\ldots\}$ proving the existence.
2. **(Uniqueness)** Assume that there exists $\varepsilon_\sigma,\varepsilon_{\sigma'}\in\{-1,1\}$ such that $\sigma\Phi=\varepsilon_\sigma\Phi$ and $\sigma\Phi=\varepsilon'_\sigma\Phi$ then $\varepsilon_\sigma\Phi=\varepsilon'_\sigma\Phi$. Now if $(x_1,\ldots,x_n)=(1,\ldots,n)$ then if $i<j$ we have $x_i-x_j=i-j\neq 0$ so that $\Phi(1,\ldots,n)=\prod_{(i,j)\in\{(i,j)\in\{1,\ldots,n\}\times\{1,\ldots,n\}|i<j\}}(x_i-x_j)\neq 0$ (see above lemma) and then from $\varepsilon_\sigma\Phi(1,\ldots,n)=\varepsilon_{\sigma'}\Phi(1,\ldots,n)$ we have by multiplying by $\Phi(1,\ldots,n)^{-1}$ that $\varepsilon_\sigma=\varepsilon_{\sigma'}$ □

Definition 10.89. *Given $n\in\mathbb{N}$ and $\sigma\in P_n$ the unique $\varepsilon_\sigma\in\{-1,1\}$ associated with σ (see the previous theorem) is called the* ***sign*** *of σ and noted $\operatorname{sign}(\sigma)$. This defines a mapping $\operatorname{sign}: P_n\to\{-1,1\}$ by $\sigma\to\operatorname{sign}(\sigma)$ where using the previous theorem $\operatorname{sign}(\sigma)=-1$ if we can write σ as a odd number of strict transpositions and $\operatorname{sign}(\sigma)$ if we can write σ as a even number of transpositions (this even works if $n=1$ and $\sigma=1_{\{0\}}$ for then $\operatorname{sign}(\sigma)=1=(-1)^0$)*

Lemma 10.90. *$\langle\{-1,1\},\cdot\rangle$ forms a abelian group*

1. **($\cdot$: $\{-1,1\}\rightarrow\{-1,1\}$ is *a* mapping)**
   a. $-1\cdot-1=1\in\{-1,1\}$
   b. $-1\cdot1=-1\in\{-1,1\}$
   c. $1\cdot-1=-1\in\{-1,1\}$
   d. $1\cdot1=1\in\{-1,1\}$
2. **(associativity)** *This follows from the fact that $\cdot$ is associative in $\mathbb{Z}$*
3. **(neutral element)** 1 *is the neutral element as* $1\cdot-1=-1\cdot1=1$ *and* $1\cdot1=1$
4. **(inverse element)** 1 *has as inverse element* 1 *and* -1 *has as inverse element* -1
5. **(commutativity)** $1\cdot-1=-1=-1\cdot1$

Theorem 10.91. *If $n\in\mathbb{N}$ then $\operatorname{sign}: P_n\rightarrow\{-1,1\}$ is a homeomorphism (see 3.13) from the group $\langle P_n,\circ\rangle$ to the multiplicative group $\langle\{-1,1\},\cdot\rangle$ or in other words:*

1. $\forall\sigma,\tau\in P_n$ *we have* $\operatorname{sign}(\sigma\circ\tau)=\operatorname{sign}(\sigma)\cdot\operatorname{sign}(\tau)$
2. $\operatorname{sign}(i_{\{1,\ldots,n\}})=1$

Proof.

1. $(\sigma\circ\tau)\Phi=\sigma(\tau\Phi)=\sigma(\operatorname{sign}(\tau)\Phi)=\operatorname{sign}(\tau)\cdot\sigma\Phi=\operatorname{sign}(\tau)\cdot(\operatorname{sign}(\sigma)\cdot\Phi)=(\operatorname{sign}(\tau)\cdot\operatorname{sign}(\sigma))\cdot\Phi\Rightarrow\operatorname{sign}(\sigma\circ\tau)=\operatorname{sign}(\sigma)\cdot\operatorname{sign}(\tau)$
2. $i_{\{1,\ldots,n\}}\Phi=\Phi=1\cdot\Phi\Rightarrow\operatorname{sign}(i_{\{1,\ldots,n\}})=1$ □

Corollary 10.92. *Given $n\in\mathbb{N}$ then if $\sigma\in P_n\Rightarrow\operatorname{sign}(\sigma)=\operatorname{sign}(\sigma^{-1})$*

Proof. As $1=\operatorname{sign}(1_{\{1,\ldots,n\}})=\operatorname{sign}(\sigma\circ\sigma^{-1})\underset{\text{previous theorem}}{=}\operatorname{sign}(\sigma)=\operatorname{sign}(\sigma^{-1})\Rightarrow\operatorname{sign}(\sigma)=\operatorname{sign}(\sigma^{-1})$ □

Definition 10.93. *If $n\in\mathbb{N}$ and $i,j\in\{1,\ldots,n\}$ then we define $(i\rightsquigarrow_n j)$ as follows*

1. **($i=j$)** *then* $(i\rightsquigarrow_n j)=1_{\{1,\ldots,n\}}$
2. **($i<j$)** *then* $(i\rightsquigarrow_n j)$ *is defined by*

$$\begin{aligned}(i\rightsquigarrow_n j)(k) &= k\text{ if }k<i\\ &= k+1\text{ if }i\leqslant k<j\\ &= i\text{ if }k=j\\ &= k\text{ if }j<k\end{aligned}$$

3. **($j<i$)** *then* $(i\rightsquigarrow_n j)$ *is defined by*

$$\begin{aligned}(i\rightsquigarrow_n j)(k) &= k\text{ if }k<j\\ &= i\text{ if }k=j\\ &= k-1\text{ if }j<k\leqslant i\\ &= k\text{ if }i<k\end{aligned}$$

Example 10.94.

1. $n=6$

$$(2 \rightsquigarrow_6 5)\begin{pmatrix}1\\2\\3\\4\\5\\6\end{pmatrix} = \begin{pmatrix}1\\3\\4\\5\\2\\6\end{pmatrix}$$

2. $n=6$

$$(5 \rightsquigarrow_6 2)\begin{pmatrix}1\\2\\3\\4\\5\\6\end{pmatrix} = \begin{pmatrix}1\\5\\2\\3\\4\\6\end{pmatrix}$$

So we interpret $(i \rightsquigarrow_n j)$ as removing i from its position an inserting it after (before) j if $i<j$ $(j<i)$.

Example 10.95. In case of $n=6$ we have

$$\begin{aligned}(2\leftrightarrow_6 5)\circ(2\leftrightarrow_6 4)\circ(2\leftrightarrow_6 3)\begin{pmatrix}1\\2\\3\\4\\5\\6\end{pmatrix} &= (2\leftrightarrow_6 5)\circ(2\leftrightarrow_6 4)\begin{pmatrix}1\\3\\2\\4\\5\\6\end{pmatrix}\\ &= (2\leftrightarrow_6 5)\begin{pmatrix}1\\3\\4\\2\\5\\6\end{pmatrix}\\ &= \begin{pmatrix}1\\3\\4\\5\\2\\6\end{pmatrix}\end{aligned}$$

$$
\begin{aligned}
&= (2 \rightsquigarrow_6 5)\begin{pmatrix}1\\2\\3\\4\\5\\6\end{pmatrix}\\
(5 \leftrightarrow_6 2)\circ(5 \leftrightarrow_6 3)\circ(5 \leftrightarrow_6 4)\begin{pmatrix}1\\2\\3\\4\\5\\6\end{pmatrix} &= (5 \leftrightarrow_6 2)\circ(5 \leftrightarrow_6 3)\begin{pmatrix}1\\2\\3\\5\\4\\6\end{pmatrix}\\
&= (5 \leftrightarrow_6 2)\begin{pmatrix}1\\2\\5\\3\\4\\6\end{pmatrix}\\
&= \begin{pmatrix}1\\5\\2\\3\\4\\6\end{pmatrix}\\
&= (5 \rightsquigarrow_6 2)\begin{pmatrix}1\\2\\3\\4\\5\\6\end{pmatrix}
\end{aligned}
$$

Theorem 10.96. *If $n \in \mathbb{N}$ then if $i, j \in \{1, \ldots, n\}$ then $\mathrm{sign}(i \rightsquigarrow_n j) = (-1)^{|i-j|}$. Further if $i \neq j$ then $(i \rightsquigarrow_n j) = ((i \leftrightarrow_n j_1) \circ \ldots \circ (i \leftrightarrow_n j_{|i-j|}))$ with $\forall k \in \{1, \ldots, |i-j|\}$ where*

$$
\begin{aligned}
j_k &= j-(k-1) \text{ if } i<j\\
&= j+(k-1) \text{ if } j<i
\end{aligned}
$$

or in other words

$$
\begin{aligned}
(i \rightsquigarrow_n j) &= ((i \leftrightarrow_n j)\circ(i \leftrightarrow_n j-1)\circ\ldots\circ(i \leftrightarrow_n i+1)) \text{ if } i<j\\
&= ((i \leftrightarrow_n j)\circ(i \leftrightarrow_n j+1)\circ\ldots\circ(i \leftrightarrow_n i-1)) \text{ if } j<i
\end{aligned}
$$

Proof. The proof is very simple but a little bit complex if you do it formally. Given $i, j \in \{1, \ldots, n\}$ we have the following cases:

1. $\boldsymbol{(i = j)}$ then $|i - j| = 0$ and $(i \rightsquigarrow_n j) = i_{\{1, \ldots, n\}}$ then $\operatorname{sign}(i \rightsquigarrow_n j) = 1 = (-1)^0 = (-1)^{|i-j|}$
2. $\boldsymbol{(i < j)}$ then we take $k = j - i$ and prove the theorem by induction on k so take $X = \{k \in \{1, \ldots\} | \text{if } i - j = k \text{ then } \operatorname{sign}(i \rightsquigarrow_n j) = (-1)^k \text{ and } (i \rightsquigarrow_n j) = ((i \leftrightarrow_n j_1) \circ \ldots \circ (i \leftrightarrow_n j_k)) \text{ where } j_k = j - (k-1)\}$ then we have :
   a. $\boldsymbol{(k = 1)}$ then $j = i + 1$ now if $l \in \{1, \ldots, n\}$ we have :
      i. $\boldsymbol{(l < i)}$ then we have
      $$\begin{array}{rcl} (i \rightsquigarrow_n i{+}1)(l) & = & l \\ & \underset{l \neq i, i+1}{=} & (i \leftrightarrow_n i+1)(l) \end{array}$$
      ii. $\boldsymbol{(i \leqslant l < i{+}1)}$ then $l = i$ and
      $$\begin{array}{rcl} (i \rightsquigarrow_n i{+}1)(l) & = & l{+}1 \\ & = & i{+}1 \\ & = & (i \leftrightarrow_n i+1)(i) \\ & \underset{l=i}{=} & (i \leftrightarrow_n i+1)(l) \end{array}$$
      iii. $\boldsymbol{(l = i{+}1)}$ then
      $$\begin{array}{rcl} (i \rightsquigarrow_n i+1)(l) & = & i \\ & \underset{l=i+1}{=} & (i \leftrightarrow_n i+1)(l) \end{array}$$
      iv. $\boldsymbol{(i + 1 < l)}$ then
      $$\begin{array}{rcl} (i \rightsquigarrow_n i+1)(l) & = & l \\ & \underset{l \neq i, i+1}{=} & (i \leftrightarrow_n i+1)(l) \end{array}$$

      So $(i \rightsquigarrow_n i+1) = (i \leftrightarrow_n i+1)$ and $\operatorname{sign}(i \rightsquigarrow_n i+1) = (-1)^1$ and thus $1 \in X$
   b. $\boldsymbol{(k \in X)}$ we must now prove that $k{+}1 \in X$. From $j - i = k + 1$ we have that $j = i + k + 1$ and $j - 1 = i + k > i$. First we prove that $(i \rightsquigarrow_n j) = (i \leftrightarrow_n j) \circ (i \rightsquigarrow_n j - 1)$ and this easel seen for if $l \in \{1, \ldots, n\}$ then we have the following cases:
      i. $\boldsymbol{(l < i)}$ then
      $$\begin{array}{rcl} (i \leftrightarrow_n j) \circ (i \rightsquigarrow_n j-1)(l) & = & (i \leftrightarrow_n j)(l) \\ & \underset{l \neq i, j}{=} & l \\ & \underset{l < i}{=} & (i \rightsquigarrow_n j)(l) \end{array}$$
      ii. $\boldsymbol{(i \leqslant l < j - 1)}$ then
      $$\begin{array}{rcl} (i \leftrightarrow_n j) \circ (i \rightsquigarrow_n j-1)(l) & = & (i \leftrightarrow_n j)(l+1) \\ & \underset{i < i+1 \leqslant l+1 < j}{=} & l+1 \\ & \underset{i \leqslant l+1 < j}{=} & (i \rightsquigarrow_n j)(l) \end{array}$$

iii. $(\boldsymbol{l=j-1})$ then

$$\begin{array}{lcl}(i\leftrightarrow_n j)\circ(i\rightsquigarrow_n j-1)(l) & = & (i\leftrightarrow_n j)(i)\\ & = & j\\ & \underset{i<j\Rightarrow i\leqslant j-1=l<j}{=} & (i\rightsquigarrow_n j)(j-1)\\ & \underset{l=j-1}{=} & (i\rightsquigarrow_n j)(l)\end{array}$$

iv. $(\boldsymbol{j-1<l})$ then we have the following cases

A. $(\boldsymbol{j=l})$ then

$$\begin{array}{lcl}(i\leftrightarrow_n j)\circ(i\rightsquigarrow_n j-1)(l) & \underset{j-1<l=j}{=} & (i\leftrightarrow_n j)(l)\\ & \underset{l=j}{=} & (i\leftrightarrow_n j)(j)\\ & = & i\\ & \underset{l=j}{=} & (i\rightsquigarrow_n j)(l)\end{array}$$

B. $(\boldsymbol{j<l})$ then

$$\begin{array}{lcl}(i\leftrightarrow_n j)\circ(i\rightsquigarrow_n j-1)(l) & \underset{j-1<j<l}{=} & (i\leftrightarrow_n j)(l)\\ & \underset{i<j<l\Rightarrow i,j\neq l}{=} & l\\ & \underset{j<l}{=} & (i\rightsquigarrow_n j)(l)\end{array}$$

proving indeed that $(i\leftrightarrow_n j)\circ(i\rightsquigarrow_n j-1)=(i\rightsquigarrow_n j)$.

Now we have

$$\begin{array}{lcl}(i\rightsquigarrow_n j) & = & (i\leftrightarrow_n j)\circ(i\rightsquigarrow_n j-1)\\ & \underset{(j-1)-i=i+(k+1)-1-i=k\wedge k\in X}{=} & (i\leftrightarrow_n j)\circ[(i\leftrightarrow_n j-1)\circ(i\leftrightarrow_n (j-1)-1)\circ\ldots\circ(i\leftrightarrow_n i+1)]\\ & = & ((i\leftrightarrow_n j)\circ(i\leftrightarrow_n j-1)\circ\ldots\circ(i\leftrightarrow_n i+1))\end{array}$$

Also $\operatorname{sign}(i\rightsquigarrow_n j)=\operatorname{sign}(i\leftrightarrow_n j)\cdot\operatorname{sign}(i\rightsquigarrow_n j-1)=(-1)\cdot(-1)^{j-1-i}=(-1)^{j-i}$, proving that $k+1\in X$ and thus $X=\{1,\ldots\}=\mathbb{N}$ proving our assertion.

3. $(\boldsymbol{j<i})$ then we take $k=i-j$ and prove the theorem by induction on k so take $X=\{k\in\{1,\ldots\}|\text{if } i-j=k \text{ then } \operatorname{sign}(i\rightsquigarrow_n j)=(-1)^k \text{ and } (i\rightsquigarrow_n j)=((i\leftrightarrow_n j_1)\circ\ldots\circ(i\leftrightarrow_n j_k)) \text{ where } j_k=j+(k-1)\}$ then we have :

a. $(\boldsymbol{k=1})$ so $j=i-1$ then we have the following cases for $l\in\{1,\ldots,n\}$

i. $(\boldsymbol{l<j})$ then

$$\begin{array}{lcl}(i\rightsquigarrow_n i-1)(l) & \underset{l<j=i-1}{=} & l\\ & \underset{l<j\leqslant i-1<i}{=} & (i\leftrightarrow_n i-1)(l)\end{array}$$

ii. $(\boldsymbol{l=j})$ then $j=i-1$

$$\begin{aligned}(i \rightsquigarrow_n i-1)(l) &= i\\ &\underset{l=j=i-1}{=} (i \leftrightarrow_n i-1)(l)\end{aligned}$$

iii. $(\boldsymbol{j<l\leqslant i})$ then $i-1=j<l\leqslant i \Rightarrow l=i$ and thus

$$\begin{aligned}(i \rightsquigarrow_n i-1)(l) &\underset{i-1<l\leqslant i}{=} l-1\\ &\underset{l=i}{=} i-1\\ &= (i \leftrightarrow_n i-1)(i)\\ &\underset{i=l}{=} (i \leftrightarrow_n i-1)(l)\end{aligned}$$

iv. $(\boldsymbol{i<l})$ then

$$\begin{aligned}(i \rightsquigarrow_n i-1)(l) &\underset{i,i-1<l}{=} l\\ &\underset{i,i-1\neq l}{=} (i \leftrightarrow_n i-1)(l)\end{aligned}$$

So $(i \rightsquigarrow_n i-1)=(i \leftrightarrow_n i-1)$ and $\operatorname{sign}(i \rightsquigarrow_n i-1)=\operatorname{sign}(i \leftrightarrow_n i-1)=(-1)=(-1)^1=(-1)^{|i-j|}$ proving that $1\in X$

b. $(\boldsymbol{k\in X})$ we have then to prove that $k+1\in X$. From $i-j=k+1$ it follows that $j=i-k-1$ and $j+1<i-k<i$. First we prove that $(i \rightsquigarrow_n j)\circ(i \rightsquigarrow_n j+1)=(i \rightsquigarrow_n j)$. So let $l\in\{1,\ldots,n\}$ then we have the following cases

i. $(\boldsymbol{l<j})$ then

$$\begin{aligned}(i \leftrightarrow_n j)\circ(i \rightsquigarrow_n j+1)(l) &\underset{l<j<j+1}{=} (i \leftrightarrow_n j)(l)\\ &\underset{l<j<i}{=} l\\ &\underset{l<j}{=} (i \rightsquigarrow_n j)(l)\end{aligned}$$

ii. $(\boldsymbol{l=j})$ then

$$\begin{aligned}(i \leftrightarrow_n j)\circ(i \rightsquigarrow_n j+1)(l) &\underset{l=j<j+1}{=} (i \leftrightarrow_n j)(l)\\ &\underset{l=j}{=} i\\ &\underset{l=j}{=} (i \rightsquigarrow_n j)(l)\end{aligned}$$

iii. $(\boldsymbol{j<l\leqslant i})$ Here we have the following sub-cases :

A. $(\boldsymbol{l=j+1})$ then

$$\begin{aligned}(i \leftrightarrow_n j)\circ(i \rightsquigarrow_n j+1)(l) &\underset{l=j+1}{=} (i \leftrightarrow_n j)(i)\\ &= j\\ &\underset{l=j+1}{=} l-1\\ &\underset{j<l\leqslant i}{=} (i \rightsquigarrow_n j)(l)\end{aligned}$$

B. $\boldsymbol{(j+1<l\leqslant i)}$ then

$$\begin{aligned}(i\leftrightarrow_n j)\circ(i\rightsquigarrow_n j+1)(l) &= (i\leftrightarrow_n j)(l-1)\\ &\underset{j<j+1\leqslant l-1<l\leqslant i}{=} l-1\\ &\underset{j<l\leqslant i}{=} (i\rightsquigarrow_n j)(l)\end{aligned}$$

iv. $\boldsymbol{(i<l)}$ then

$$\begin{aligned}(i\leftrightarrow_n j)\circ(i\rightsquigarrow_n j+1)(l) &= (i\leftrightarrow_n j)(l)\\ &\underset{j<i<l}{=} l\\ &\underset{i<l}{=} (i\rightsquigarrow_n j)(l)\end{aligned}$$

So we have that

$$\begin{aligned}(i\rightsquigarrow_n j) &= (i\leftrightarrow_n j)\circ(i\rightsquigarrow_n j+1)\\ &\underset{\text{induction and } i-(j+1)=k}{=} (i\leftrightarrow_n j)\circ[(i\leftrightarrow_n j+1)\circ(i\leftrightarrow_n j+2)\circ\ldots\circ(i\leftrightarrow_n i-1)]\\ &= ((i\leftrightarrow_n j)\circ(i\leftrightarrow_n j+1)\circ\ldots\circ(i\leftrightarrow_n i-1))\end{aligned}$$

also $\operatorname{sign}(i\rightsquigarrow_n j)=\operatorname{sign}(i\leftrightarrow_n j)\cdot\operatorname{sign}(i\rightsquigarrow_n j)\underset{k\in X}{=}(-1)\cdot(-1)^k=(-1)^{k+1}$ proving that $k+1\in X$ and by mathematical induction that $X=\{1,\ldots\}=\mathbb{N}$ □

Theorem 10.97. *Given $\tau\in P_n$ define then $T_\tau: P_n\rightarrow P_n$ by $\sigma\rightarrow T_\tau(\sigma)=\tau\circ\sigma$ is a bijection*

Proof.

1. **(surjectivity)** If $\sigma\in P_n$ take then $\sigma'=\tau^{-1}\circ\sigma\in P_n$ then $T_\tau(\sigma')=\tau\circ(\tau^{-1}\circ\sigma)=(\tau\circ\tau^{-1})\circ\sigma=1_{\{1,\ldots,n\}}\circ\sigma=\sigma$ so we have found a σ' such that $T_\tau(\sigma')=\sigma$
2. **(injectivity)** If $T_\tau(\sigma)=T_\tau(\sigma')\Rightarrow\tau\circ\sigma=\tau\circ\sigma'\Rightarrow\tau^{-1}\circ(\tau\circ\sigma)=\tau^{-1}\circ(\tau\circ\sigma')=(\tau^{-1}\circ\tau)\circ\sigma=(\tau^{-1}\circ\tau)\circ\sigma'=1_{\{1,\ldots,n\}}\circ\sigma=1_{\{1,\ldots,n\}}\circ\sigma'=\sigma=\sigma'$ □

10.3 Finite product of sets

Definition 10.98. *Let $n\in\mathbb{N}$, A a set then we define A^n as $A^n=\prod_{i\in\{1,\ldots,n\}}A_i$ where $\{A_i\}_{i\in\{1,\ldots,n\}}$ where $\forall i\in\{1,\ldots,n\}$ we have $A_i=A$.*

Theorem 10.99. *Let $n\in\mathbb{N}$ and A a set then we have that $A^n=A^{\{1,\ldots,n\}}$*

Proof. See 2.82 □

We can of course define the power of a set by cartesian products

Definition 10.100. *Let $n\in\mathbb{N}$, A a set then $A^{(\times)n}$ is defined recursively by:*

$n=1$. *then $A^{(\times)1}=A$*

$n>1$. *then $A^{\times(n)}=(A^{(\times)(n-1)})\times A$*

***Theorem* 10.101.** *Let $n\in\mathbb{N}$, A a set then $A^{(\times)n}=\bigotimes_{i\in\{1,\ldots,n\}}A_i$ where $\{A_i\}_{i\in\{1,\ldots,n\}}$ is such that $A_i=A$*

Proof. *This is trivially proved by recursion so let $S_A=\left\{n\in\mathbb{N}|A^{(\times)n}=\bigotimes_{i\in\{1,\ldots,n\}}A_i \text{ where } \{A_i\}_{i\in\{1,\ldots,n\}} \text{ is such that } A_i=A\right\}$ then we have*

$1\in S$. *$A\underset{\text{definition}}{=}A^{(\times)1}\underset{\text{definition}}{=}\bigotimes_{i\in\{1,\ldots,1\}}A_i$ where $\{A_i\}_{i\in\{1,\ldots,1\}}$ is such that $A_1=A$ proving that $1\in S$.*

$n\in S\Rightarrow n+1\in S$. *Let $\{A_i\}_{i\in\{1,\ldots,n+1\}}$ such that $A_i=A$ then we have $A^{(\times)(n+1)}=(A^{(\times)n})\times A\underset{n\in S}{=}\left(\bigotimes_{i\in\{1,\ldots,n\}}A_i\right)\times A=\left(\bigotimes_{i\in\{1,\ldots,n\}}A_i\right)\times A_{n+1}=\bigotimes_{i\in\{1,\ldots,n+1\}}A_i$ proving $n+1\in S$* □

Let's now compare the two definitions of the finite power of a set.

Theorem 10.102. *Let $n\in\mathbb{N}$ and A as set then there exists a bijection $\mathcal{P}\colon A^{(\times)n}\to A^n$ so that*

1. *If $B\subseteq A$ then $\mathcal{P}(B^{(\times)n})=B^n$*
2. *$\forall[x_1,\ldots,x_n]\in B^{(\times)n}$ we have $\mathcal{P}([x_1,\ldots,x_n])=(x_1,\ldots,x_n)$*

Proof. This follows from the previous theorem and 5.88 □

Theorem 10.103. *Let $n\in\mathbb{N}$ and $\{A_i\}_{i\in\{1,\ldots,n\}}$ be such that $\forall i\in\{1,\ldots,n\}$ we have A_i is finite then $\bigotimes_{i\in\{1,\ldots,n\}}A_i$ and $\prod_{i\in\{1,\ldots,n\}}A_i$ are finite. Furthermore we have that $\#\left(\bigotimes_{i\in\{1,\ldots,n\}}A_i\right)=\#\left(\prod_{i\in\{1,\ldots,n\}}A_i\right)=\prod_{i=1}^{1}\#(A_i)$*

Proof. We prove this by mathematical induction so let $S=\left\{n\in\mathbb{N}|\{A_i\}_{i\in\{1,\ldots,n\}} \text{ a finite family of finite sets then } \bigotimes_{i\in\{1,\ldots,n\}}A_i \text{ is finite with } \#\left(\bigotimes_{i\in\{1,\ldots,n\}}A_i\right)=\prod_{i\in\{1,\ldots,n\}}\#(A_i)\right\}$ then we have

$1\in S$. If $\{A_i\}_{i\in\{1,\ldots,n\}}$ is a finite family of finite sets then $\bigotimes_{i\in\{1,\ldots,1\}}A_i=A_i$ which is finite and $\#\left(\bigotimes_{i\in\{1,\ldots,1\}}A_i\right)=\#(A_1)=\prod_{i=1}^{1}\#(A_i)$ so that $1\in S$

$n\in S\Rightarrow n+1\in S$. Let $\{A_i\}_{i\in\{1,\ldots,n+1\}}$ then $\bigotimes_{i\in\{1,\ldots,n+1\}}A_i=\left(\bigotimes_{i\in\{1,\ldots,n\}}A_i\right)\times A_{n+1}$ which, as $\bigotimes_{i\in\{1,\ldots,n\}}A_n$ is finite $(n\in S)$ and A_{n+1} is finite, is finite by 5.44. Also $\#\left(\bigotimes_{i\in\{1,\ldots,n+1\}}A_i\right)=\#\left(\left(\bigotimes_{i\in\{1,\ldots,n\}}A_i\right)\times A_{n+1}\right)\underset{5.44}{=}\#\left(\bigotimes_{i\in\{1,\ldots,n\}}A_i\right)\cdot\#(A_{n+1})\underset{n\in S}{=}\left(\prod_{i=1}^{n}\#(A_i)\right)\cdot\#(A_{n+1})=\prod_{i=1}^{n+1}\#(A_i)$ proving that $n+1\in S$

As $\bigotimes_{i\in\{1,\ldots,n\}}A_i$ and $\prod_{i\in\{1,\ldots,n\}}A_i$ are bijective we have proved the rest of our theorem. □

A easy consequence of the above theorem is the following theorem.

Theorem 10.104. *If $n \in \mathbb{N}$ and X is a finite set then X^n and $X^{(\times)n}$ is finite and $\#(X^{(\times)n}) = \#(X^n) = \#(X)^n$*

Proof. See the previous theorem. □

10.4 Vector Spaces

10.4.1 Definition and examples

Definition 10.105. *A vector space $\langle V, \oplus, \odot \rangle$ over a field $\langle F, +, . \rangle$ is a abelian group $\langle V, \oplus \rangle$ together with a map $\odot : F \times V \to V$ (here we note $\odot(x, y)$ as $x \odot y$ and call this scalar multiplication of a scalar and a vector) such that*

1. *$\forall \alpha \in F, \forall x, y \in V$ we have $\alpha \odot (x \oplus y) = (\alpha \odot x) \oplus (\alpha \odot y)$*
2. *$\forall \alpha, \beta \in F, \forall x \in V$ we have $(\alpha + \beta) \odot x = (\alpha \odot x) \oplus (\beta \odot y)$*
3. *$\forall \alpha, \beta \in F, \forall x \in V$ we have $(\alpha . \beta) \odot x = \alpha \odot (\beta \odot x)$*
4. *$\forall x \in V$ we have $1 \odot x = x$ (1 is the neutral element in $\langle F, \cdot \rangle$)*

*Elements of V are called **vectors**.*

Theorem 10.106. *Let $\langle V, \oplus, \odot \rangle$ be a vector space over a field $\langle F, +, . \rangle$ then we have*

1. *$\forall x \in V$ we have $O = 0 \odot x$ (O is the neutral element in $\langle V, \oplus \rangle$ and 0 is the neutral element in $\langle F, + \rangle$*
2. *$\forall x \in V$ we have $(-1) \odot x = -x$ where -1 is the inverse element of 1 in $\langle F, + \rangle$ and $-x$ is the inverse element of x in $\langle V, \oplus \rangle$*
3. *$\forall \alpha \in F$ we have $\alpha \odot O = O$*

Proof.

1. $O = 0 \odot x \oplus (-(0 \odot x)) \underset{0=0+0}{=} (0 + 0) \odot x \oplus (-(0 \odot x)) = (0 \odot x + 0 \odot x) \oplus (-(0 \odot x)) = 0 \odot x \oplus (0 \odot x \oplus (-(0 \odot x))) = 0 \odot x \oplus O = 0 \odot x$
2. $x \oplus ((-1) \odot x) = (1 \odot x) \oplus ((-1) \odot x) = (1 + (-1)) \odot x = 0 \odot x \underset{(1)}{=} O \Rightarrow (-1) \odot x = -x$
3. $\alpha \odot O \underset{(1)}{=} \alpha \odot (0 \odot O) \underset{3.25}{=} (\alpha \cdot 0) \odot O = 0 \odot O = O$ □

Theorem 10.107. *Let $\langle V, \oplus, \odot \rangle$ be a vector space over a field $\langle F, +, . \rangle$ then if $\alpha \in F \setminus \{0\}$ (0 is the neutral element in $\langle F, + \rangle$) we have $\forall x \in V\ \alpha \odot x = O \Rightarrow x = O$ (O the neutral element in $\langle V, \oplus \rangle$)*

Proof. As $\alpha \neq 0$ we have that α^{-1} exists. $O = \alpha^{-1} \odot O \underset{\alpha \odot x = O}{=} \alpha^{-1} \odot (\alpha \odot x) = (\alpha^{-1}.\alpha) \odot x = 1 \odot x = x$ □

Definition 10.108. *If $\langle V, \oplus, \odot \rangle$ is a vector space over a field $\langle F, +, \cdot \rangle$ and $\emptyset \neq W \subseteq V$ is a non-empty set such that $\forall \alpha, \beta \in F$ and $\forall x, y \in W$ we have that $\alpha \odot x \oplus \beta \odot y \in W$ then W is called a subspace of V*

Theorem 10.109. *If $\langle V, \oplus, \odot \rangle$ is a vector space over a field $\langle F, +, \cdot \rangle$ and $\emptyset \neq W \subseteq V$ is a subspace of V then $\langle W, \oplus_{|W \times W}, \odot_{|F \times W} \rangle$ is a vector space over a field $\langle F, +, \cdot \rangle$. Note that in general we write the subspace as $\langle W, \oplus, \odot \rangle$ and we assume it is now that we use here the restrictions to W of the vector operations.*

Proof. First we must prove that $\langle W, \oplus_{|W \times W} \rangle$ is a abelian group

1. $\oplus_{|W \times W}: W \times W \rightarrow W$ is a function, for this to be true we must have $\forall x$, $y \in W$ that $x \oplus_{|W \times W} y = x \oplus y \in W$. Now $x \oplus y = (1 \odot x) \oplus (1 \odot y) \in W$ as W is a subspace.
2. **(associativity)** If $x < y, z \in W$ then $x \oplus_{|W \times W} (y \oplus_{|W \times W} z) = x \oplus (y \oplus z) = (x \oplus y) \oplus z = (x \circ_{|W \times W} y) \oplus_{|W \times W} z$
3. **(neutral element)** If O is the neutral element of $\langle V, \oplus \rangle$ then as $\emptyset \neq W$ there exists a $w \in W$ but then $O = O \oplus O \underset{10.106}{=} 0 \odot x \oplus 0 \odot x \in W$ as W is a subspace. Now if $x \in W$ then $x \oplus_{|W \times W} O = x \oplus O = x = O \oplus x = O \oplus_{|W \times W} x$
4. **(inverse element)** If $x \in W$ then $-x \underset{10.106}{=} (-1) \odot x \oplus O \in W$ (see (3) and W is a subspace) and $x \oplus_{|W \times W} (-x) = x \oplus (-x) = O = (-x) \oplus x = (-x) \oplus_{|W \times W} x$
5. **(commutativity)** If $x, y \in W$ then $x \oplus_{|W \times W} y = x \oplus y = y \oplus x = y \oplus_{|W \times W} x$

Next we prove the remaining axioms of a vector space

1. $\odot_{|F \times W}: F \times W \rightarrow W$ is a function, for this to be true we must have that $\forall \alpha \in F, \forall x \in W$ that $\alpha \odot_{|F \times W} x = \alpha \odot x \in W$. But $\alpha \odot x = \alpha \odot x + O = \alpha \odot x + 0 \odot x \in W$
2. If $\alpha \in F$ and $x, y \in W$ then $\alpha \odot_{|W \times W} (x \oplus_{|W \times W} y) = \alpha \odot (x \oplus y) = \alpha \odot x \oplus \alpha \odot y = \alpha \odot_{|W \times W} x \oplus_{|W \times W} \beta \odot_{|W \times W} y$
3. If $\alpha, \beta \in F$ and $x \in W$ then $(\alpha + \beta) \odot_{|W \times W} x = (\alpha + \beta) \odot x = \alpha \odot x \oplus \beta \odot x = \alpha \odot_{|W \times W} x \oplus_{|W \times W} \beta \odot_{|W \times W} x$
4. If $\alpha, \beta \in F$ and $x \in W$ then $(\alpha \cdot \beta) \odot_{|W \times W} x = (\alpha \cdot \beta) \odot x = \alpha \odot (\beta \odot x) = \alpha \odot_{|W \times W} (\beta \odot_{|W} x)$
5. If $x \in W$ then $1 \odot_{|F \times W} x = 1 \odot x = x$ □

Let's now look at some examples of vector spaces.

First every field is a vector space over itself.

Theorem 10.110. *If $\langle F, +, \cdot\rangle$ is a field then $\langle F, +, \cdot\rangle$ is a vector space over the field $\langle F, +, \cdot\rangle$*

Proof. As F is a field we have that (see 3.29) $\langle F, +\rangle$ is a abelian group and further from the axioms of a field we have :

1. $\forall \alpha \in F, \forall x, y \in F$ we have $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$
2. $\forall \alpha, \beta \in F, \forall x \in F$ we have $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$
3. $\forall \alpha, \beta \in F, \forall x \in F$ we have $(\alpha \cdot \beta) \cdot x = \alpha \cdot (\beta \cdot x)$
4. $\forall x \in F$ we have $1 \cdot x = x$ □

As $\langle \mathbb{R}, +, \cdot\rangle$ is a field (see 8.25) we have thus

Corollary 10.111. *$\langle \mathbb{R}. +, \cdot\rangle$ is a vector space over itself. $\langle \mathbb{C}, +, \cdot\rangle$ is a vector space over itself*

We can also make $\langle \mathbb{C}, +, \cdot\rangle$ a vector space over $\langle \mathbb{R}, +, \cdot\rangle$

Theorem 10.112. *$\langle \mathbb{C}, +, \cdot\rangle$ is a vector space over $\langle \mathbb{R}, +, \cdot\rangle$ where the scalar multiplication $\cdot: \mathbb{R} \times \mathbb{C} \rightarrow \mathbb{C}$ is defined to be $(\alpha, (x, y)) \rightarrow (\alpha \cdot x, \alpha \cdot y)$*

Proof. As $\langle \mathbb{C}, +, \cdot\rangle$ is a field we have that $\langle \mathbb{C}, +\rangle$ is a abelian group. We have then

1. $\forall \alpha \in \mathbb{R}, \forall (x, y), (x', y') \in \mathbb{C}$ we have $\alpha \cdot ((x, y) + (x', y')) = \alpha \cdot (x + x', y + y') = (\alpha \cdot (x + x'), \alpha \cdot (y + y')) = (\alpha \cdot x + \alpha \cdot x', \alpha \cdot y + \alpha \cdot y') = (\alpha \cdot x, \alpha \cdot y) + (\alpha \cdot x', \alpha \cdot y') = \alpha \cdot (x, y) + \alpha \cdot (x', y')$
2. $\forall \alpha, \beta \in \mathbb{R}, \forall (x, y) \in \mathbb{C}$ we have $(\alpha + \beta) \cdot (x, y) = ((\alpha + \beta) \cdot x, (\alpha + \beta) \cdot y) = (\alpha \cdot x + \beta \cdot x, \alpha \cdot y + \beta \cdot y) = (\alpha \cdot x, \alpha \cdot y) + (\beta \cdot x, \beta \cdot y) = \alpha \cdot (x, y) + \beta \cdot (x, y)$
3. $\forall \alpha, \beta \in \mathbb{R}, \forall (x, y) \in \mathbb{C}$ we have $(\alpha \cdot \beta) \cdot (x, y) = ((\alpha \cdot \beta) \cdot x, (\alpha \cdot \beta) \cdot y) = (\alpha \cdot (\beta \cdot x), \alpha \cdot (\beta \cdot y)) = \alpha \cdot (\beta \cdot x, \beta \cdot y) = \alpha \cdot (\beta \cdot (x, y))$
4. $\forall (x, y) \in F$ we have $1 \cdot (x, y) = (1 \cdot x, 1 \cdot y) = (x, y)$ □

Let's see now how we can create a vector space based on a existing vector space.

Theorem 10.113. *Let $\langle V, \oplus, \odot\rangle$ be a vector space over a field $\langle F, +, \cdot\rangle$, X a set then $V^X = \{f | f: X \rightarrow V \text{ is a function}\}$ (the set of functions from X to V (see 2.69)). Define then $\boxplus: V^X \times V^X \rightarrow V^X$ by $(f, g) \rightarrow f + g$ where $(f + g)(x) = f(x) + g(x)$ and $\boxdot: F \times V^X \rightarrow V^X$ by $(\alpha, f) \rightarrow \alpha \boxdot f$ where $(\alpha \boxdot f)(x) = \alpha \odot f(x)$. Then $\langle V^X, \boxplus, \boxdot\rangle$ forms a vector space over $\langle F, +, \cdot\rangle$ we note this vector space as $M(X, V)$.*

Proof. First we prove that $\langle V^X, \boxplus\rangle$ forms a abelian group with units $0: X \rightarrow V$ where $0(x) = 0$ for every x in X. So let $f, g, h \in V^X$ then we have:

1. **(associativity)** $\forall x \in X$ we have $f \boxplus (g \boxplus h)(x) = f(x) \oplus (g \boxplus h)(x) = f(x) \oplus (g(x) \oplus h(x)) \underset{\langle V, \oplus, \odot\rangle \text{ is a vector space}}{=} (f(x) \oplus g(x)) \oplus h(x) = (f \boxplus g)(x) \oplus h(x) = ((f \boxplus g) \boxplus)(x) \Rightarrow f \boxplus (g \boxplus h) = (f \boxplus g) \boxplus h$

2. **(commutativity)** $\forall x \in X$ we have $(f \boxplus g)(x) = f(x) \oplus g(x) \underset{\langle V, \oplus, \odot \rangle \text{ is a vector space}}{=} g(x) \oplus f(x) = (g \boxplus f)(x) \Rightarrow f \boxplus g = g \boxplus f$
3. **(neutral element)** $\forall x \in X$ we have $(f \boxplus 0)(x) = f(x) \oplus 0(x) = f(x) \oplus 0 = f(x) \Rightarrow f \boxplus 0 = f$ and $0 \boxplus f = f$ follows then from commutativity.
4. **(inverse element)** Define $-f: X \rightarrow V$ by $x \rightarrow -f(x)$ then we have $\forall x \in X$ that $(f \boxplus (-f))(x) = f(x) \oplus (-f(x)) \underset{\langle V, \oplus, \odot \rangle \text{ is a vector space}}{=} 0 \Rightarrow f \boxplus (-f) = 0$ and $(-f) \boxplus f = 0$ follows from commutativity.

For the remaining vector space axioms note that:

1. If $\alpha \in F$ and $f, g \in V^X$ then $\forall x \in X$ we have $(\alpha \boxdot (f \boxplus g))(x) = \alpha \odot (f \boxplus g)(x) = \alpha \odot (f(x) \oplus g(x)) \underset{\langle V, \oplus, \odot \rangle \text{ is a vector space}}{=} \alpha \odot f(x) \oplus \alpha \odot g(x) = (\alpha \boxdot f)(x) \oplus (\alpha \boxdot g)(x) = ((\alpha \boxdot f) \boxplus (\alpha \boxdot g))(x) \Rightarrow \alpha \boxdot (f \boxplus g) = \alpha \boxdot f \boxplus \alpha \boxdot g$
2. If $\alpha, \beta \in F$ and $g \in V^X$ then $\forall x \in X$ we have $((\alpha + \beta) \boxdot f)(x) = (\alpha + \beta) \odot f(x) \underset{\langle V, \oplus, \odot \rangle \text{ is a vector space}}{=} \alpha \odot f(x) \oplus \beta \odot f(x) = (\alpha \boxdot f)(x) \oplus (\beta \boxdot f)(x) = ((\alpha \boxdot f) \boxplus (\beta \boxdot f))(x) \Rightarrow (\alpha + \beta) \boxdot f = (\alpha \boxdot f) \boxplus (\beta \boxdot f)$
3. If $\alpha, \beta \in F$ and $f \in V^X$ then $\forall x \in X$ we have $((\alpha \cdot \beta) \boxdot f)(x) = (\alpha \cdot \beta) \odot f(x) \underset{\langle V, \oplus, \odot \rangle \text{ is a vector space}}{=} \alpha \cdot (\beta \cdot f(x)) = \alpha \cdot ((\beta \boxdot f)(x)) = (\alpha \boxdot (\beta \boxdot f))(x) \Rightarrow (\alpha \cdot \beta) \boxdot f = \alpha \boxdot (\beta \boxdot f)$
4. If $1 \in F$ and $f \in V^x$ then $\forall x \in V$ we have $(1 \boxdot f)(x) = 1 \odot f(x) \underset{\langle V, \oplus, \odot \rangle \text{ is a vector space}}{=} f(x) \Rightarrow 1 \boxdot f = f$ $\square$

Note 10.114. To avoid excessive use of different use of different operator notations we use from now on only the symbols $+, \cdot$ instead of $\oplus, \boxplus, \odot$ and $\boxdot$ and assume that care full inspection of the context where a symbol is used is enough to find to which space (field ...) a operator belongs. This is similar to our use of $0, 1$.

If $\langle V, +, \cdot \rangle$ is a vector space over $\langle F, +, \cdot \rangle$ and if $n \in \mathbb{N}$ then by 10.99 we have that $V^n = V^{\{1, \ldots, n\}}$ and by applying the above theorem we have automatically that the following corollary is true.

Corollary 10.115. *If $\langle V, +, \cdot \rangle$ is a vector space over $\langle F, +, \cdot \rangle$ and $n \in \mathbb{N}$ then $\langle V^n, +, \cdot \rangle$ is a vector space over $\langle F, +, \cdot \rangle$ where $+: V^n \times V^n \rightarrow V^n$ is defined by $(x, y) \rightarrow x + y$ where $\forall i \in \{1, \ldots, n\}$ we have $(x + y)_i = (x + y)(i) = x(i) + y(i) = x_i + y_i$ and $\cdot: F \times V^n \rightarrow V^n$ is defined by $(\alpha, x) \rightarrow \alpha \cdot x$ where $\forall i \in \{1, \ldots, n\}$ we have $(\alpha \cdot x)_i = (\alpha \cdot x)(i) = \alpha \cdot x(i) = \alpha \cdot x_i$*

This with the fact that a field forms a vector space over itself gives us the following corollaries (where we have assumed the above meaning of $+, \cdot$ in F^n

Corollary 10.116. *If $\langle F, +, \cdot \rangle$ is a field and $n \in \mathbb{N}$ then $\langle F^n, +, \cdot \rangle$ forms a vector space over the field $\langle F, +, \cdot \rangle$*

Corollary 10.117. *If $n \in \mathbb{N}$ then*

1. *$\langle \mathbb{R}^n, +, \cdot \rangle$ forms a vector space over $\langle \mathbb{R}, +, \cdot \rangle$*
2. *$\langle \mathbb{C}^n, +, \cdot \rangle$ forms a vector space over $\langle \mathbb{C}, +, \cdot \rangle$*
3. *$\langle \mathbb{C}^n, +, \cdot \rangle$ forms a vector space over $\langle \mathbb{R}, +, \cdot \rangle$*

Theorem 10.118. *If $\langle V_i, +_i, \cdot_i \rangle_{i \in I}$ is a family of vecto spaces over a field $\langle F, +, \cdot \rangle$ then if we define:*

1. *$+: \prod_{i \in I} V_i \times \prod_{i \in I} V_i \rightarrow \prod_{i \in I} V_i$ is defined by $(x, y) \rightarrow x + y$ where $x + y$: $I \rightarrow \bigcup_{i \in I} V_i$ is defined by $(x + y)(i) = x(i) +_i y(i) = x_i +_i y_i$ (see 3.15)*
2. *$\cdot: F \times \prod_{i \in I} V_i \rightarrow \prod_{i \in I} V_i$ is defined by $(\alpha, x) \rightarrow \alpha \cdot x$ where $\alpha \cdot x: I \rightarrow \bigcup_{i \in I} V_i$ is defined by $(\alpha \cdot x) = \alpha \cdot_i x(i) = \alpha \cdot_i x_i$*

then we have that $\langle \prod_{i \in I} V_i, +, \cdot \rangle$ is a vector space over $\langle F, +, \cdot \rangle$

Proof. From 3.15 it follows that $\langle \prod_{i \in I} V_i, + \rangle$ is a abelian group. Next if $\alpha \in F$ and $x \in V_i$ we have by the fact that $\langle V_i, +_i, \cdot_i \rangle$ is a group that $\alpha \cdot_i x(i) \in V_i$ that $\alpha \cdot x: I \rightarrow \bigcup_{i \in I} V_i$ is a element of $\prod_{i \in I} V_i$ and thus that $\cdot: F \times \prod_{i \in I} V_i \rightarrow \prod_{i \in I} V_i$ is indeed a function. Now that we have proved that (1) and (2) are well defined we prove the rest of the vector space axioms.

1. If $\alpha \in F$ and $x, y \in \prod_{i \in I} V_i$ then $\forall i \in I$ we have $(\alpha \cdot (x + y))(i) = \alpha \cdot_i (x + y)(i) = \alpha \cdot_i (x(i) +_i y(i)) \underset{\langle V_i, +_i, \cdot_i \rangle \text{ is a vector space}}{=} \alpha \cdot_i x(i) +_i \alpha \cdot_i y(i) = (\alpha \cdot x)(i) +_i (\alpha \cdot y)(i) = (\alpha \cdot x + \alpha \cdot y)(i)$ and thus $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$
2. If $\alpha, \beta \in F$ and $x \in \prod_{i \in I} V_i$ then $\forall i \in I$ we have $((\alpha + \beta) \cdot x)(i) = (\alpha + \beta) \cdot_i x(i) \underset{\langle V_i, +_i, \cdot_i \rangle}{=} \alpha \cdot_i x(i) +_i \beta \cdot_i x(i) = (\alpha \cdot x)(i) +_i (\beta \cdot x)(i) = (\alpha \cdot x + \beta \cdot y)(i)$ and thus $(\alpha + \beta) \cdot x$
3. If $\alpha, \beta \in F$ and $x \in \prod_{i \in I} V_i$ then $\forall i \in I$ we have $((\alpha \cdot \beta) \cdot x)(i) = (\alpha \cdot \beta) \cdot_i x(i) \underset{\langle V_i, +_i, \cdot_i \rangle \text{ is a vector space}}{=} \alpha \cdot_i (\beta \cdot_i x(i)) = \alpha \cdot_i (\beta \cdot x)(i) = (\alpha \cdot (\beta \cdot x))(i)$ and thus $(\alpha \cdot \beta) \cdot x = \alpha \cdot (\beta \cdot x)$
4. If 1 is the unit in F and $x \in \prod_{i \in I} V_i$ then $\forall i \in I$ we have $(1 \cdot x)(i) = 1 \cdot_i x(i) \underset{\langle V_i, +_i, \cdot_i \rangle \text{ is a vector space}}{=} x(i)$ and thus $1 \cdot x = x$ □

Note 10.119. To avoid having to use excessive notation we do not use anymore the notation: $\langle V, +, \cdot \rangle$ is a vector space over a field $\langle F, +, \cdot \rangle$

Instead we just say that V is a vector space over the field F and assume that $+, \cdot$ are the additive and multiplicative operators in V, F (using context to find out which is which). Also we assume that the multiplicative units in V is 1 and the neutral element in V and F is 0 (again context will tell us which is which).

Definition 10.120. *If V is a vector space over a field F and $A, B \subseteq V$, $G \subseteq F$, $x \in V$ and $\alpha \in F$ then we have the following definitions*

1. *$x + A = \{x + y | y \in A\} = \{y + x | y \in A\} = A + x$*

2. $A+B=\{x+y|x\in A\wedge x\in B\}$
3. $\alpha\cdot A=\{\alpha\cdot x|x\in A\}$
4. $G\cdot A=\{\alpha\cdot x|\alpha\in G\wedge x\in A\}$
5. $A-B=\{x-y|x\in A\wedge y\in B\}$

Theorem 10.121. *If V is a vector space over a field F then we have*

1. $\forall A\subseteq V, x\in V$ *we have* $y\in x+A\Leftrightarrow y-x\in A$
2. $\forall A,B\subseteq V, x\in V$ *we have* $x+(A\bigcup B)=(x+A)\bigcup(x+B)$
3. $\forall A,B\subseteq V, x\in V$ *we have* $x+(A\bigcap B)=(x+A)\bigcap(x+B)$
4. $\forall A\subseteq V, x,y\in V$ *we have* $x+(y+A)=(x+y)+A$

Proof.

1. *If* $y\in x+A$ *there exists a* $z\in A$ *such that* $y=x+z\Rightarrow y-x=z\in A$. *If* $y-x\in A\Rightarrow y=x+(y-x)\in x+A$.
2. $y\in x+(A\bigcup B)\Leftrightarrow y-x\in A\bigcup B\Leftrightarrow(y-x\in A)\vee(y-x\in B)\Leftrightarrow(y\in x+A)\vee(y\in x+B)\Leftrightarrow y\in(x+A)\bigcup(x+B)$
3. $y\in x+(A\bigcap B)\Leftrightarrow y-x\in A\bigcap B\Leftrightarrow(y-x\in A)\wedge(y-x\in B)\Leftrightarrow(y\in x+A)\wedge(y\in x+B)\Leftrightarrow y\in(x+A)\bigcap(x+B)$
4. $z\in(x+(y+A))\Leftrightarrow z-x\in y+A\Leftrightarrow(z-x)-y\in A\Leftrightarrow(z-(x+y))\in A\Leftrightarrow z\in(x+y)+A$ □

10.4.2 Linear dependency

Lemma 10.122. *If V is a vector space over a field F, $\alpha\in F$, $n\in\mathbb{N}_0$ and $\{x_i\}_{i\in\{0,\ldots,n\}}$ a family in V then $\sum_{i=0}^{n}\alpha\cdot x_i=\alpha\cdot\sum_{i=0}^{n}x_i$*

Proof. We prove this by mathematical induction so let $X=\{n\in\mathbb{N}_0|\text{if } \{x_i\}_{i\in\{0,\ldots,n\}} \text{ is } a \text{ family of elements in } V \text{ then } \sum_{i=0}^{n}\alpha\cdot x_i=\alpha\cdot\sum_{i=0}^{n}\}$ we have then:

1. For $n=0$ we have $\sum_{i=0}^{0}\alpha\cdot x_i=\alpha\cdot x_0=\alpha\cdot\sum_{i=0}^{0}x_i$ so $0\in X$
2. If $n\in X$ and if $\{x_i\}_{i\in\{0,\ldots,n+1\}}$ is a family of elements in V then $\sum_{i=0}^{n+1}\alpha\cdot x_i=(\sum_{i=0}^{n}\alpha\cdot x_i)+(\alpha\cdot x_{n+1})\underset{n\in X}{=}\alpha\cdot(\sum_{i=0}^{n}\alpha\cdot x_i)+\alpha\cdot x_{n+1}=\alpha\cdot(\sum_{i=0}^{n}x_i+x_{n+1})=\alpha\cdot\sum_{i=1}^{n+1}x_i$ so $n+1\in X$

Using mathematical induction we have then $X=\mathbb{N}_0$ proving our theorem. □

Lemma 10.123. *If V is a vector space over a field F, $\{x_i\}_{i\in I}$ a family of sets with finite support and $\alpha\in F$ then $\{\alpha\cdot x_i\}_{i\in I}$ is a family of sets with finite support (so $\sum_{i\in I}\alpha\cdot x_i$ is defined) and $\sum_{i\in I}\alpha\cdot x_i=\alpha\cdot\sum_{i\in I}x_i$*

Proof. If $i \in \mathrm{support}(\{\alpha \cdot x_i\}_{i \in I})$ then $\alpha \cdot x_i \neq 0$, now if $x_i = 0$ then by 10.106 we have $\alpha \cdot x_i = 0$ contradicting $\alpha \cdot x_i \neq 0$ so we must have $x_i \neq 0$ and thus $i \in \mathrm{support}(\{x_i\}_{i \in I})$ or in other words $\mathrm{support}(\{\alpha \cdot x_i\}_{i \in I}) \subseteq \mathrm{support}(\{x_i\}_{i \in I})$. So as $\mathrm{support}(\{\alpha \cdot x_i\}_{i \in I})$ is finite we have by 5.36 that $\mathrm{support}(\{\alpha \cdot x_i\}_{i \in I})$ is finite.

Using the definition of the generalized sum (see 10.37 and 10.52) we have $\sum_{i \in I} \alpha \cdot x_i = \sum_{i=0}^{n-1} \alpha \cdot x_{b(i)}$ where $b: S_n \to \mathrm{support}(\{x_i\}_{i \in I})$ is a bijection. By the previous lemma we have then that $\sum_{i=0}^{n-1} \alpha \cdot x_{b(i)} = \alpha \cdot \sum_{i=0}^{n-1} x_{b(i)} = \alpha \cdot \sum_{i \in I} x_i$ and thus $\sum_{i \in I} \alpha \cdot x_i = \alpha \cdot \sum_{i \in I} x_i$. □

Lemma 10.124. *If V is a vector space over a field F, $v \in V$ and $\{\alpha_i\}_{i \in \{0,\ldots,n\}}$ a family in F then $\sum_{i=0}^{n} \alpha_i \cdot v = (\sum_{i=0}^{n} \alpha_i) \cdot v$*

Proof. We prove this by induction so let $S = \{n \in \mathbb{N}_0 | \text{if } \{\alpha_i\}_{i \in \{0,\ldots,n\}} \text{ is a family in } F \text{ and } v \in V \text{ then } \sum_{i=0}^{n} \alpha_i \cdot v = (\sum_{i=0}^{n} \alpha_i) \cdot v\}$. So we have:

1. If $n = 0$ then $\sum_{i=0}^{0} \alpha_i \cdot v = \alpha_0 \cdot v = (\sum_{i=0}^{0} \alpha_i) \cdot v$
2. If $n \in S$ then if $\{\alpha_i\}_{i \in \{0,\ldots,n+1\}}$ is a family in F and $v \in V$ then $\sum_{i=0}^{n+1} \alpha_i \cdot v = (\sum_{i=0}^{n} \alpha_i \cdot v) + \alpha_{n+1} \cdot v \underset{n \in S}{=} (\sum_{i=0}^{n} \alpha_i) \cdot v + \alpha_{n+1} \cdot v = (\sum_{i=0}^{n} \alpha_i + \alpha_{n+1}) \cdot v = (\sum_{i=0}^{n+1} \alpha_i) \cdot v$ and thus $n + 1 \in S$

Using mathematical induction we have then $S = \mathbb{N}_0$ proving our theorem. □

Lemma 10.125. *If V is a vector space over a field F, $v \in V$ and $\{\alpha_i\}_{i \in I}$ a family in F with finite support then $\{\alpha_i \cdot v\}_{i \in I}$ has finite support and $\sum_{i \in I} \alpha_i \cdot v = (\sum_{i \in I} \alpha_i) \cdot v$*

Proof. If $i \in \mathrm{support}(\{\alpha_i \cdot v\}_{i \in I}) \Rightarrow \alpha_i \cdot v \neq 0 \Rightarrow \alpha_i \neq 0 \Rightarrow i \in \mathrm{support}(\{\alpha_i\}_{i \in I}) \Rightarrow \mathrm{support}(\{\alpha_i \cdot v\}_{i \in I}) \subseteq \mathrm{support}(\{\alpha_i\}_{i \in I})$ and thus $\mathrm{support}(\{\alpha_i \cdot v\}_{i \in I})$ is finite. If $b: S_n \to \mathrm{support}(\{\alpha_i\}_{i \in I})$ then using the definition of the generalized sum (see 10.37 and 10.52) we have $\sum_{i \in I} \alpha_i \cdot v = \sum_{i=0}^{n-1} \alpha_i \cdot v \underset{\text{previous lemma}}{=} (\sum_{i=0}^{n-1} \alpha_i) \cdot v = (\sum_{i \in I} \alpha_i) \cdot v$ □

Theorem 10.126. *If V is a vector space over a field F, $\{v_i\}_{i \in I}$ a family in V and $\{\alpha_i\}_{i \in I}$ a family in F with finite support then $\{\alpha_i \cdot v_i\}_{i \in I}$ is a family in V with finite support and $\mathrm{support}(\{\alpha_i \cdot v_i\}_{i \in I}) \subseteq \mathrm{support}(\{\alpha_i\}_{i \in I})$.*

Proof. If $i \in \mathrm{support}(\{\alpha_i \cdot v_i\}_{i \in I}) \Rightarrow \alpha_i \cdot v_i \neq 0 \Rightarrow \alpha_i \neq 0$ [otherwise $\alpha_i \cdot v_i = 0$]$\Rightarrow i \in \mathrm{support}(\{\alpha_i\}_{i \in I}) \Rightarrow \mathrm{support}(\{\alpha_i \cdot v_i\}_{i \in I}) \subseteq \mathrm{support}(\{\alpha_i\}_{i \in I})$. The theorem follows then from 5.36 and the fact that $\mathrm{support}(\{\alpha_i\}_{i \in I})$ is finite. □

The above theorem ensures that the next definition make sense.

Definition 10.127. *Let V be a vector space over a field F then if $\{v_i\}_{i \in I}$ is a family in V and $v \in V$ then v is a **linear combination of $\{v_i\}_{\in I}$** if there exists a family $\{\alpha_i\}_{i \in I}$ with finite support such that $v = \sum_{i \in I} \alpha_i \cdot v_i$ (sum is defined as $\{\alpha_i \cdot v_i\}_{i \in I}$ has finite support by the previous theorem). If $W \subseteq V$ then a **linear combination of W** is defined to be the **linear combination** of the family **$\{w\}_{w \in W}$ (see 2.107)**.*

Definition 10.128. *Let V be a vector space over a field F, $W \subseteq V$ in V then* ***$\mathcal{S}(W)$ the linear span of W*** *is defined by $\mathcal{S}(W) = \{v \in V | \exists \{\alpha_w\}_{w \in W}$ with finite support such that $v = \sum_{w \in W} \alpha_w \cdot w\}$. If $\{v_i\}_{i \in I}$ is a family of elements in V then $\mathcal{S}(\{v_i\}_{i \in I})$ is defined by $\mathcal{S}(\{v_i\}_{i \in I}) \underset{\text{definition}}{=} \mathcal{S}(\{v_i | i \in I\})$ (see 2.104).*

Definition 10.129. *Let V be a vector space over a field F, $A \subseteq V$ then if $W \subseteq V$ we say that W spans A if $\mathcal{S}(W) = A$*

Theorem 10.130. *Let V be a vector space over a field F, $W \subseteq V$ then $\mathcal{S}(W)$ is a subspace of V. If $W \neq \emptyset$ then $\mathcal{S}(W)$ is a vector space. (and thus by 10.109 $\mathcal{S}(W)$ is a sub space).*

Proof. Let $\alpha, \beta \in F$ and $x, y \in \mathcal{S}(W)$ then there exists $\{\alpha_w\}_{w \in W}, \{\beta_w\}_{w \in W}$ such that $x = \sum_{w \in W} \alpha_w \cdot w$, $y = \sum_{w \in W} \beta_w \cdot w$. So $\alpha \cdot x + \beta \cdot y = \alpha \cdot \sum_{w \in W} \alpha_w \cdot w + \beta \cdot \sum_{w \in W} \beta_w \cdot w \underset{10.123}{=} \sum_{i \in w} \alpha \cdot (\alpha_w \cdot w) + \sum_{w \in W} \beta \cdot (\beta_w \cdot w) \underset{10.53}{=} \sum_{w \in W} (\alpha \cdot (\alpha_w \cdot w) + \beta \cdot (\beta_w \cdot w)) = \sum_{w \in W} (\alpha \cdot \alpha_w + \beta \cdot \beta_w) \cdot w \in \mathcal{S}(W)$.

If $W \neq \emptyset$ then there exists a $w \in W$ and then $0 = 0 \cdot w \in \mathcal{S}(W)$, using 10.109 we have then that $\mathcal{S}(W)$ is a vector space. □

Theorem 10.131. *Let V be a vector space over a field F and $W \subseteq V$ a subspace of V then*

1. *If $\{w_i\}_{i \in \{0, \ldots, n\}}$ is a family in W then $\sum_{i=0}^{n} w_i \in W$*
2. *If $\{w_i\}_{i \in I}$ is a family in W with finite support then $\sum_{i \in I} w_i \in W$*

Proof.

1. We proceed by induction, let $S = \{n \in \mathbb{N} | \text{If } \{w_i\}_{i \in \{0, \ldots, n\}} \text{ is a family in } W \text{ then } \sum_{i=0}^{n} w_i \in W\}$, then:

   a. $\sum_{0}^{0} w_i = w_0 \in W \Rightarrow 0 \in S$

   b. If $n \in S$ take then $\{w_i\}_{i \in \{0, \ldots, n+1\}}$ in W then $\sum_{i=0}^{n+1} w_i = (\sum_{i=0}^{n} w_i) + w_{n+1} \in W$ [As $\sum_{i=0}^{n} w_i \in W$ (because $n \in S$), $w_{n+1} \in W$ and W is a subspace] so $n + 1 \in S$

   Using mathematical induction we have then $S = \mathbb{N}_0$ proving (1).

2. If $\{w_i\}_{i \in I}$ is a family with finite support then by 10.37 there exists a bijection $b: \{0, \ldots, n-1\} \rightarrow \text{support}(\{w_i\}_{i \in I})$ such that $\sum_{i \in I} w_i = \sum_{i=0}^{n-1} w_{b(i)} \in W$ (see (1)) □

It is now easy to write a equivalent definition of $\mathcal{S}(W)$ using finite families only

Theorem 10.132. *Let V be a vector space over a field F and $W \subseteq V$ then $\mathcal{S}(W) = \{x \in V | \exists \{\alpha_i\}_{i \in \{0, \ldots, n-1\}}$ a finite family in F and $\exists \{v_i\}_{i \in \{0, \ldots, n-1\}}$ a finite family in V such that $x = \sum_{i \in \{0, \ldots, n-1\}} \alpha_i \cdot v_i\}$ (note we use $n-1$ to indicate that $\{0, \ldots, n-1\}$ can be the empty set (so that the sum is 0)).*

Proof. Let $\mathcal{V} = \{x \in V | \exists \{\alpha_i\}_{i \in \{1, \ldots, n\}}$ a finite family in F and $\exists \{v_i\}_{i \in \{1, \ldots, n\}}$ a finite family in V such that $x = \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot v_i\}$. If now $x \in \mathcal{V}$ then there exists a $\{\alpha_i\}_{i \in \{0, \ldots, n\}}$ and $\{v_i\}_{i \in \{0, \ldots, n\}}$ such that $x = \sum_{i \in \{0, \ldots, n\}} \alpha_i \cdot v_i = \sum_{i=0}^{n} \alpha_i \cdot v_i$ which as $\forall i \in \{1, \ldots, n\}$ we have $\alpha_i \cdot v_i \in \mathcal{S}(W)$ [as $\boldsymbol{S}(W)$ is a vector space and $W \subseteq \mathcal{S}(W)$] means by the above theorem that $x \in \mathcal{S}(W)$ so we have proved that

$$\mathcal{V} \subseteq \mathcal{S}(W) \tag{10.11}$$

If $x \in \mathcal{S}(W)$ then there exists a family $\{\alpha_w\}_{w \in W}$ with finite support such that $x = \sum_{w \in W} \alpha_w \cdot w$. As $\operatorname{support}(\{\alpha_w \cdot w\}_{w \in W})$ is finite there exists a bijection $b: \{0, \ldots, n-1\} \rightarrow \operatorname{support}(\{\alpha_w \cdot w\}_{w \in \{0, \ldots, n\}})$. If we define then the finite families $\{\alpha_{b(i)}\}_{i \in \{0, \ldots, n-1\}}$, $\{b(i)\}_{i \in \{0, \ldots, n-1\}}$ then we have that $x = \sum_{w \in W} \alpha_w \cdot w \underset{10.37}{=} \sum_{i=0}^{n-1} \alpha_{b(i)} \cdot b(i) = \sum_{i \in \{0, \ldots, n-1\}} \alpha_{b(i)} \cdot b(i) \in \mathcal{V}$ proving that

$$\mathcal{S}(W) \subseteq \mathcal{V} \tag{10.12}$$

Using 10.11 and 10.12 we have $\mathcal{V} = \mathcal{S}(W)$ proving our theorem. □

Theorem 10.133. *Let V be a vector space over a field F and $W \subseteq V$ then we have $W \subseteq \mathcal{S}(W)$*

Proof. If $w_0 \in W$ define then $\{\alpha_w\}_{w \in W}$ a family in F by $\alpha_w = \left\{ \begin{smallmatrix} 1 \text{ if } w = w_0 \\ 0 \text{ if } w \in W \setminus \{w_0\} \end{smallmatrix} \right.$, this family clearly has finite support and $\sum_{w \in W} \alpha_w \cdot w = \sum_{i \in W \setminus \{w_0\}} \alpha_w \cdot w + \sum_{w \in \{w_0\}} \alpha_w \cdot w = \sum_{w \in W \setminus \{w_0\}} 0 \cdot w + 1 \cdot w_0 = w_0$ so $w_0 \in \mathcal{S}(W)$ □

Theorem 10.134. *Let V be a vector space over the field F and $W \subseteq V$ a subspace of V then $\mathcal{S}(W) = W$*

Proof. By 10.133 we have that $W \subseteq \mathcal{S}(W)$. Now if $v \in \mathcal{S}(W)$ then $\exists \{\alpha_w\}_{w \in W}$ in F with finite support such that $v = \sum_{w \in W} \alpha_w \cdot w$. Using 10.131 and the fact that $\alpha_w \cdot w \in W$ we have $\sum_{w \in W} \alpha_w \cdot w \in W \Rightarrow v \in W \Rightarrow \mathcal{S}(W) \subseteq W$ and thus $W = \mathcal{S}(W)$. □

Corollary 10.135. *If V is a vector space over a field F and $W \subseteq V$ then $\mathcal{S}(W) = \mathcal{S}(\mathcal{S}(W))$*

Proof. If $W = \emptyset$ then $\mathcal{S}(W) = \emptyset$ and thus $\mathcal{S}(\mathcal{S}(W)) = \emptyset = \mathcal{S}(W)$. If $W \neq \emptyset$ then use the above theorem and the fact that $\mathcal{S}(W)$ is a subspace (see 10.130) is $W \neq \emptyset$. □

Theorem 10.136. *Let V be a vector space over a field F, $W_1, W_2 \subseteq V$ with $W_1 \subseteq W_2$ then $\mathcal{S}(W_1) \subseteq \mathcal{S}(W_2)$*

Proof. We have the following cases to consider:

1. $(\boldsymbol{W_1 = W_2})$ then trivially $\mathcal{S}(W_1) = \mathcal{S}(W_2) \Rightarrow \mathcal{S}(W_1) \subseteq \mathcal{S}(W_2)$

2. $(\boldsymbol{W_1 \subset W_2})$ if $v \in \mathcal{S}(W_1)$ then there exists a $\{\alpha_w\}_{w\in W_1}$ in F with finite support such that $v = \sum_{w\in W_1} \alpha_w \cdot w$. Define then $\{\beta_w\}_{w\in W_2}$ by $\beta_w = \begin{cases} \alpha_w \text{ if } w \in W_1 \\ 0 \text{ if } w \in W_2 \setminus W_1 \end{cases}$ then $\operatorname{support}(\{\alpha_w\}_{w\in W_1}) = \operatorname{support}(\{\beta_w\}_{w\in W_2})$ and thus $\{\beta_w\}_{w\in W_2}$ has finite support, further $W_2 \ni \sum_{w\in W_2} \beta_w \cdot w = \sum_{w\in W_2\setminus W_1} \beta_w \cdot w + \sum_{w\in W_1} \beta_w \cdot w = \sum_{w\in W_2\setminus W_1} 0 \cdot w + \sum_{w\in W_1} \alpha_w \cdot w = v$ so $\mathcal{S}(W_1) \subseteq \mathcal{S}(W_2)$ □

Definition 10.137. *Let V be a vector space over a field F and $\{v_i\}_{i\in I}$ a family of elements in V then $\{v_i\}_{i\in I}$ is **linear independent** if $\forall \{\alpha_i\}_{i\in I}$ in F with finite support such that $\sum_{i\in I} \alpha_i \cdot v_i = 0$ we have that $\forall i \in I$ $\alpha_i = 0$ (in other words* $\operatorname{support}(\{a_i\}_{i\in I}) = \emptyset$. *If $\{v_i\}_{i\in I}$ is not linear independent it is **linear dependent**.*

Note that we can also define linear dependency and linear dependency for a set based on the fact that we can given as set W define the family $\{w\}_{w\in W}$ by self indexing (see 2.101)

Definition 10.138. *If V is a vector space over a field F and $W \subseteq V$ then W is linear independent (or linear dependent) if $\{w\}_{w\in W}$ is linear independent (or linear dependent)*

The following theorem gives a relation between linear independence of a family and its associated set.

Theorem 10.139. *If V is a vector space over a field F and $\{v_i\}_{i\in I}$ a family in V. If $b: J \to I$ is a bijection then $\{v_i\}_{i\in I}$ is linear dependent (or linear independent) then $\{v_{b_j}\}_{j\in J}$ is linear dependent (or linear independent).*

Proof. First we observe that if $\{\alpha_i\}_{i\in I}$ is a family in F with finite support we can define a family $\{\alpha'_j\}_{j\in J}$ by $\alpha'_j = \alpha_{b(j)}$ and thus $\sum_{j\in J} \alpha'_j \cdot v_{b_j} = \sum_{j\in J} \alpha_{b(j)} \cdot v_{b(j)} \underset{10.44}{=} \sum_{i\in I} \alpha_i \cdot v_i$. Also if $\{\beta_j\}_{j\in J}$ is a family in F with finite support we can define a family $\{\beta''_i\}_{i\in I}$ by $\beta''_i = \beta_{b^{-1}(i)}$ so that $\beta''_{b(i)} = \beta_{b^{-1}(b(i))} = \beta_i$ we have then that $\sum_{j\in J} \beta_j \cdot v_{b_i} = \sum_{j\in J} \beta''_{b(i)} \cdot v_{b(i)} = \sum_{i\in I} \beta''_i \cdot v_i$

1. If $\{v_i\}_{i\in I}$ is linear dependent then there exists a $\{\alpha_i\}_{i\in I}$ such that $\exists i_0 \in I \vdash \alpha_{i_0} \neq 0$ and $0 = \sum_{i\in I} \alpha_i \cdot v_i = \sum_{j\in J} \alpha'_j \cdot v_{b_j}$ where for $j_0 = b(i_0)$ we have $\alpha'_{j_0} = \alpha_{b(i_0)} \neq 0$ proving that $\{v_{b_j}\}_{j\in J}$ is linear dependent.
2. If now $\{v_i\}_{i\in I}$ is linear independent and $\{\beta_j\}_{j\in J}$ a family in F such that $0 = \sum_{j\in J} \beta_j \cdot v_{b_j} = \sum_{i\in I} \beta''_i \cdot v_i$ and by linear Independence we have then that $\forall i \in I$ we have $\beta''_i = 0$. Then $\forall j \in J$ we have $\beta_j = \beta''_{b^{-1}(i)} = 0$ proving linear Independence of $\{v_{b_j}\}_{j\in J}$ □

Theorem 10.140. *Let V be a vector space over a field F, $\{v_i\}_{i\in I}$ a family in V with $\exists i, j \in I, i \neq j$ with $v_i = v_j$ then $\{v_i\}_{i\in I}$ is linear dependent.*

Proof. If $i, j \in I$ with $i \neq j$ and $v_i = v_j$ define then $\{\alpha_i\}_{i\in I}$ by $\alpha_k = \begin{cases} 1 \text{ if } k = i \\ -1 \text{ if } k = j \\ 0 \text{ if } k \in I \setminus \{i, j\} \end{cases}$ then $\{\alpha_i\}_{i\in I}$ has finite support and is not all zero. We have then $\sum_{k\in I} \alpha_k \cdot v_k = \sum_{k\in I\setminus\{i,j\}} \alpha_k \cdot v_k + \sum_{k\in\{i\}} \alpha_k \cdot v_k + \sum_{k\in\{j\}} \alpha_k \cdot v_k = \sum_{k\in I\setminus\{i,j\}} 0 \cdot v + \alpha_i \cdot v_i + \alpha_j \cdot v_j = 1 \cdot v_i + (-1) \cdot v_j = v_i + (-v_j) \underset{v_i = v_j}{=} 0.$ □

Corollary 10.141. *Let V be a vector space over a field F and $\{v_i\}_{i\in I}$ a linear independent family then if $v_i = v_j$ we have that $i = j$ (so $\{v_i\}_{i\in I}: I \rightarrow \{v_i | i \in I\}$ forms a bijection).*

Proof. If $v_i = v_j$ with $i \neq j$ then $\{v_i\}_{i\in I}$ is linear dependent contradicting the linear Independence. □

Theorem 10.142. *If V is a vector space over over a field F and $\{v_i\}_{i\in I}$ is a linear independent family in V then $\{v_i | i \in I\}$ is also linear independent.*

Proof. Define $V = \{v_i | i \in I\}$ then using the above corollary we have that $v: I \rightarrow V$ defined by $i \rightarrow v_i$ is a bijection. Assume now that $\{\alpha_w\}_{w\in V}$ is a family in F with finite support such that $\sum_{w\in V} \alpha_w \cdot w = 0$. Define then $\{\beta_i\}_{i\in I}$ by $\beta_i = \alpha_{v(i)}$,then $\text{support}(\{\beta_i\}_{i\in I}) = v^{-1}(\text{support}(\{\alpha_w\}_{w\in V}))$ is finite, and we have $\sum_{i\in I} \beta_i \cdot v_i = \sum_{i\in I} \alpha_{v(i)} \cdot v(i) \underset{10.44}{=} \sum_{w\in V} \alpha_w \cdot w = 0$ so by linear Independence of $\{v_i\}_{i\in I}$ we have that $\forall i \in I$ we have $\forall i \in I$ that $\beta_i = 0$. Finally $\forall w \in V$ we have $v^{-1}(w) \in I$ and $0 = \beta_{v^{-1}(w)} = \alpha_{v(v^{-1}(w))} = \alpha_w \Rightarrow \forall w \in V$ we have $\alpha_w = 0$ proving that $V = \{v_i | i \in I\}$ is linear independent. □

Note that the opposite is not true, for example if $0 \neq v \in V$ and we define $\{v_i\}_{i\in\{1,2\}}$ by $v_1 = v = v_2$ then $\{v_i | i \in \{1, 2\}\} = \{v\}$ is a linear independent but $\{v_i\}_{i\in\{1,2\}}$ is not as $v_1 = v_2$. However if there are no duplicates in a family then the opposite is true.

Theorem 10.143. *If V is a vector space over a field F and $\{v_i\}_{i\in I}$ a family such that $\forall i, j \in I$ with $v_i = v_j$ we have $i = j$. Then we have*

1. *$\{v_i\}_{i\in I}$ is linear independent if and only if $\{v_i | i \in I\}$ is linear independent*
2. *$\{v_i\}_{i\in I}$ is linear dependent if and only if $\{v_i | i \in I\}$ is linear dependent*

Proof. Define $V = \{v_i | i \in I\}$ then $v: I \rightarrow V$ defined by $i \rightarrow v(i) = v_i$ (the actual definition of a family) is a bijection.

1. Assume that $\{v_i\}_{i\in I}$ is linear independent then using 10.142 we have that $\{v_i | i \in I\}$ is linear independent.
2. Assume that $\{v_i | i \in I\}$ is linear independent. Then if $\{\alpha_i\}_{i\in I}$ is a family in F with finite support such that $0 = \sum_{i\in I} \alpha_i \cdot v_i$ define then $\{\beta_w\}_{w\in V}$ by $\beta_w = \alpha_{v^{-1}(w)}$ then $\text{support}(\{\beta_w\}_{w\in V}) = v(\text{support}(\{\alpha_i\}_{i\in I}))$ is finite, $\beta_{v(i)} = \alpha_{v^{-1}(v(i))} = \alpha_i$ and $0 = \sum_{i\in I} \alpha_i \cdot v_i = \sum_{i\in I} \beta_{v(i)} \cdot v(i) \underset{10.44}{=} \sum_{w\in V} \beta_w \cdot w$ and by linear independence of $V = \{v_i | i \in I\}$ we have $\forall w \in V$ that $\beta_w = 0$. If now $i \in I$ then $v(i) \in W \Rightarrow \alpha_i = \beta_{v(i)} = 0$ so we have $\forall i \in I$ that α_i proving that $\{v_i\}_{i\in I}$ is linear independent.
3. Assume that $\{v_i\}_{i\in I}$ is linear dependent then if $\{v_i | i \in I\}$ would be linear independent we would have by (2) the contradiction that $\{v_i\}_{i\in I}$ would be linear independent. So we must have that $\{v_i | i \in I\}$ is linear dependent.
4. Assume that $\{v_i | i \in I\}$ is linear dependent then if $\{v_i\}_{i\in I}$ would be linear independent we would have by (1) the contradiction that $\{v_i | i \in I\}$ would be linear independent. So we must have that $\{v\}_{i\in I}$ is linear dependent. □

Theorem 10.144. *Let V be a vector space over a field F then $\{v_i\}_{i\in I}$ in V is linear dependent if and only if $\exists i\in I$ and a $\{\alpha_k\}_{k\in I\setminus\{i\}}$ with finite support such that $v_i=\sum_{k\in I\setminus\{i\}}\alpha_k\cdot v_k$*

Proof.

If $\{v_i\}_{i\in I}$ is linear dependent then there exists a family $\{\alpha_i\}_{i\in I}$ in F with finite support such that $\sum_{i\in I}\alpha_i\cdot v_i=0$ and $\exists i\in I$ such that $\alpha_i\neq 0$. We have then that $0=\sum_{j\in I\setminus\{i\}}\alpha_i\cdot v_i+\sum_{j\in\{i\}}\alpha_j\cdot v_j=\sum_{j\in I\setminus\{i\}}\alpha_j\cdot v_j+\alpha_i\cdot v_i$ and thus as $\alpha_i\neq 0$ we have

$$v_i=(-\alpha_i^{-1})\cdot\sum_{j\in I\setminus\{i\}}\alpha_j\cdot v_j \tag{10.13}$$

If we take now $\{\beta_j\}_{j\in I\setminus\{i\}}$ by $\beta_j=-\alpha_i^{-1}\cdot\alpha_j$ which has trivially finite support (as $\{\alpha_i\}_{i\in I}$ has finite support) then we have $\sum_{j\in I\setminus\{i\}}\beta_i\cdot v_i=\sum_{j\in I\setminus\{i\}}(-\alpha_i^{-1}\cdot\alpha_j)\cdot v_j=\sum_{j\in I\setminus\{i\}}(-\alpha_i^{-1})\cdot(\alpha_j\cdot v_j)=(-\alpha_i^{-1})\cdot\sum_{j\in I\setminus\{i\}}\alpha_j\cdot v_j\underset{10.13}{=}v_i\Rightarrow\sum_{j\in I\setminus\{i\}}\alpha_j\cdot v_j=v_i$.

If there exists a $i\in I$ and a $\{\alpha_k\}_{k\in I\setminus\{i\}}$ with finite support such that $v_i=\sum_{k\in I\setminus\{i\}}\alpha_k\cdot v_k$ then

$$-v_i+\sum_{k\in I\setminus\{i\}}\alpha_k\cdot v_k=0 \tag{10.14}$$

take now $\{\beta_k\}_{k\in I}$ by $\beta_k=\left\{\begin{smallmatrix}\alpha_k \text{ if } k\in I\setminus\{i\}\\ -1 \text{ if } k=i\end{smallmatrix}\right.$ then we have $\sum_{k\in I}\beta_k\cdot v_k=\sum_{k\in I\setminus\{i\}}\beta_k\cdot v_k+\sum_{k\in\{i\}}\beta_k\cdot v_k=\sum_{k\in I\setminus\{i\}}\alpha_k\cdot v_k+(-v_i)\underset{10.14}{=}0$ so we have $\sum_{k\in I}\beta_k\cdot v_k=0$ and $\beta_i=-1\neq 0$ proving that $\{v_i\}_{i\in I}$ is linear dependent. □

Theorem 10.145. *Let V be a vector space over a field F, $\{v_i\}_{i\in I}$ a family in V with $0\in\{v_i|i\in I\}$ then $\{v_i\}_{i\in I}$ is linear dependent.*

Proof. As $0\in\{v_i|i\in I\}$ there exists a $i\in I$ such that $0=v_i$, if we take now $\{\alpha_k\}_{k\in I\setminus\{i\}}$ defined by $\alpha_k=0$ if $k\in I\setminus\{i\}$. Then $\sum_{k\in I\setminus\{i\}}\alpha_k\cdot v_k=\sum_{k\in I\setminus\{i\}}0\cdot v_k=0=v_i$. Using the above theorem we have then that $\{v_i\}_{i\in I}$ is linear dependent. □

Definition 10.146. *If S,I,J are sets and $\{x_i\}_{i\in I}$, $\{y_j\}_{j\in J}$ are families in S then we say that $\{x_i\}_{i\in I}\preccurlyeq\{y_j\}_{j\in J}$ if $I\subseteq J$ and $\forall i\in I$ we have $x_i=y_i$. We say that the $\{y_i\}_{i\in J}$ is a* ***extension*** *of $\{x_i\}_{i\in I}$. Note that if $X,Y\in S$ with $X\subseteq Y$ then trivially we have $\{x\}_{x\in X}\preccurlyeq\{y\}_{y\in Y}$*

Theorem 10.147. *Let V be a vector space over a field F, $\{v_i\}_{i\in I}$ a linear dependent family in V and $\{w_i\}_{i\in J}$ a family in V such that $\{v_i\}_{i\in I}\preccurlyeq\{w_i\}_{i\in J}$. In particular this means that if $W_1,W_2\subseteq V$ is such that W_1 is linear dependent and $W_1\subseteq W_2$ then W_2 is linear dependent.*

Proof. As $\{v_i\}_{i\in I}$ is linear dependent there exists a family $\{\alpha_i\}_{i\in I}$ in F with finite support not all zero such that $\sum_{i\in I}\alpha_i\cdot v_i$. Then as $I\subseteq J$ we can take then $\{\beta_i\}_{i\in J}$ defined by $\beta_i=\left\{\begin{smallmatrix}\alpha_i \text{ if } i\in I\\ 0 \text{ if } i\in J\setminus I\end{smallmatrix}\right.$ which has trivially finite support and is not all zero. We have then that $\sum_{i\in J}\beta_i\cdot w_i=\sum_{i\in J\setminus I}\beta_i\cdot w_i+\sum_{i\in I}\beta_i\cdot w_i=\sum_{i\in J\setminus I}0\cdot w_i+\sum_{i\in I}\alpha_i\cdot w_i\underset{\forall i\in I\vDash v_i=w_i}{=}\sum_{i\in I}\alpha_i\cdot v_i=0$ proving that $\{w_i\}_{i\in J}$ is linear dependent. □

Definition 10.148. *Let V be a vector space over a field F then a family $\{v_i\}_{i\in I}$ in V is a basis if*

1. $\mathcal{S}(\{v_i|i\in I\})=V$
2. $\{v_i\}_{i\in I}$ *is linear independent*

Theorem 10.149. *Let V be a vector space over a field F then if $\{v_i\}_{i\in I}$ is a basis of V and $b: J\to I$ is a bijection we have that $\{v_{b_i}\}_{i\in I}$ is a basis of V.*

Proof. Using 10.139 we have that $\{v_{b_i}\}_{i\in J}$ is linear independent. If $v\in\{v_{b(i)}|i\in J\}$ then $\exists i\in J$ such that $v=v_{b_i}\in\{v_i|i\in I\}\Rightarrow\{v_{b_i}|i\in J\}\subseteq\{v_i|i\in I\}$ also if $v\in\{v_i|i\in I\}$ then $\exists i\in I$ such that $v=v_i=v_{b(b^{-1}(i))}\in\{v_{b_j}|j\in J\}\Rightarrow\{v_i|i\in I\}\subseteq\{v_{b_j}|j\in J\}$ giving that $\{v_i|i\in I\}=\{v_{b_j}|j\in J\}$ and thus $V=\mathcal{S}(\{v_i|i\in I\})=\mathcal{S}(\{v_{b_j}|j\in J\})$. □

Theorem 10.150. *If V is a non trivial vector space (so $V\neq\{0\}$) with a basis $\{v_i\}_{i\in I}$ then $I\neq\emptyset$*

Proof. If V is non trivial then $\exists v\in V$ such that $v\neq 0$. Then there must exists a $\{\alpha_i\}_{i\in I}$ with finite support such that $v=\sum_{i\in I}\alpha_i\cdot v_i$, if now $\operatorname{support}(\{\alpha_i\}_{i\in I})=\emptyset$ then we would have $0=\sum_{i\in I}\alpha_i\cdot v_i=v\neq 0$ a contradiction. So there must exists a $i\in I$ such that $\alpha_i\neq 0\Rightarrow I\neq\emptyset$. □

Theorem 10.151. *Let V be a vector space over a field F and $\{v_i\}_{i\in I}$ a family in V then $\{v_i\}_{i\in I}$ is a basis of V if and only if $\forall v\in V$ there exists a **unique** family $\{\alpha_i\}_{i\in I}$ in F with finite support such that $v=\sum_{i\in I}\alpha_i\cdot v_i$.*

Proof.

If $\{v_i\}_{i\in I}$ is a basis then $\{v_i\}_{i\in I}$ is linear independent and $\mathcal{S}(\{v_i|i\in I\})=V$ so if $v\in V$ then there exists a $\{\alpha_i\}_{i\in I}$ with finite support such that $v=\sum_{i\in I}\alpha_i\cdot v_i$. Now to prove uniqueness let's assume that there exists another $\{\beta_i\}_{i\in I}$ with finite support so that $v=\sum_{i\in I}\beta_i\cdot v_i$ then we have $0=v+(-1)\cdot v=\sum_{i\in I}\alpha_i\cdot v_i+(-1)\cdot\sum_{i\in I}\beta_i\cdot v_i=\sum_{i\in I}\alpha_i\cdot v_i+\sum_{i\in I}(-\beta_i)\cdot v_i=\sum_{i\in I}(\alpha_i-\beta_i)\cdot v_i \underset{\{v_i\}_{i\in I}\text{is independent}}{\Rightarrow}\forall i\in I$ we have $\alpha_i-\beta_i=0\Rightarrow\alpha_i=\beta_i$ and thus we find that $\{\alpha_i\}_{i\in I}=\{\beta_i\}_{i\in I}$.

Assume now that $\{v_i\}_{i\in I}$ is a family in V such that $\forall v\in V$ there exists a unique $\{\alpha_i\}_{i\in I}$ with finite support so that $v=\sum_{i\in I}\alpha_i\cdot v_i$. Then clearly $\mathcal{S}(\{v_i\}_{i\in I})=V$. If now $\{\alpha_i\}_{i\in I}$ is a family in F with finite support so that $0=\sum_{i\in I}\alpha_i\cdot v_i$ if we define now $\{\beta_i\}_{i\in I}$ where $\forall i\in I$ we have β_i then we have also $0=\sum_{i\in I}\beta_i\cdot v_i$, form uniqueness it follows that $\forall i\in I$ we have $\alpha_i=\beta_i=0$ and thus we have proved linear Independence of $\{v_i\}_{i\in I}$. □

Notation 10.152. *Let V be a vector space over a field F with basis $\{e_i\}_{i\in I}$ then given $v\in V$ there exists a unique $\{\alpha_i\}_{i\in I}$ in F such that $v=\sum_{i\in I}\alpha_i\cdot e_i$ we note then $\forall i\in I$ then unique element $\alpha_i\in F$ as v_i and call this the i-the coordinate of v in the basis $\{e_i\}_{i\in I}$. So using this notation we have $v=\sum_{i\in\{1,\ldots,n\}}v_i\cdot e_i$*

Definition 10.153. *If S,I are sets then if $\{x_i\}_{i\in I}$ is a family of elements in S and $T\subseteq S$ then we say that $\{x_i\}_{i\in I}\subseteq T$ if $\{x_i|i\in I\}\subseteq T$*

The next theorem says that every linear independent family that is part of a subset that **spans** the vector space can be extented to a basis of the vector space.

Theorem 10.154. *Let V be a non trivial vector space over a field F, $W \subseteq V$ is such that $\mathcal{S}(W) = V$, R a linear independent set (see 10.138) such that $R \subseteq W$ then there exist a set B such that $\{b\}_{b \in B}$ forms a basis of V and $R \subseteq B \subseteq W$*

Proof. Using the notation of the theorem define the following set of sets

$$\mathcal{A}(R, W) = \{X \subseteq V | R \subseteq X \subseteq W \text{ and } X \text{ is linear independent}\} \tag{10.15}$$

then using 2.133 we have that

$$\langle \mathcal{A}(R, W), \subseteq \rangle \text{ is a partial ordered set} \tag{10.16}$$

Assume now that $\mathcal{C} \subseteq \mathcal{A}(R, W)$ is a non-empty chain (see 2.144) then if we define $B_{\mathcal{C}} = \bigcup_{X \in \mathcal{C}} X$ we have

$$\forall X \in \mathcal{C} \text{ we have that } X \subseteq B_{\mathcal{C}} = \bigcup_{X \in \mathcal{C}} X \tag{10.17}$$

Now if $X \in \mathcal{C} \subseteq \mathcal{A}(R, W)$ then by 10.15 we have $R \subseteq X \subseteq W$ and thus using 10.17 we have

$$R \subseteq B_{\mathcal{C}} \subseteq W \tag{10.18}$$

We prove now the following assertion by mathematical induction

$$\text{If } A \subseteq B_{\mathcal{C}} \text{ is a finite subset of } B_{\mathcal{C}} \text{ then there exists a } C \in \mathcal{C} \text{ with } A \subseteq \mathcal{C} \tag{10.19}$$

Proof. Let $S_{\mathcal{C}} = \{n \in \mathbb{N}_0 | \text{if } A \subseteq B_{\mathcal{C}} \text{ is such that } S_n \approx A \text{ then there exists a } C \in \mathcal{C} \text{ such that } A \subseteq C\}$ then we have:

1. If $n = 0$ then if $A \subseteq B_{\mathcal{C}}$ is such that $\emptyset = S_0 \approx A \Rightarrow A = \emptyset$ as $\mathcal{C}$ is not empty there exists a $C \in \mathcal{C}$ and then $A = \emptyset \subseteq C \Rightarrow 0 \in S_{\mathcal{C}}$
2. Assume that $n \in S_{\mathcal{C}}$ if now $A \subseteq B_{\mathcal{C}}$ is such that $S_{n+1} \approx A$ then there exists a bijection $b: \{0, \ldots, n\} = S_{n+1} \rightarrow A$, form then the bijection $b_{|S_n}$: $S_n = \{0, \ldots, n-1\} \rightarrow A \setminus \{b(n)\}$, so as then $S_n \approx A \setminus \{b(n)\} \subseteq B_{\mathcal{C}}$ we have as $n \in S_{\mathcal{C}}$ that there exists a $C' \in \mathcal{C}$ such that $A \setminus \{b(n)\} \subseteq C'$. Also as $b(n) \in A \subseteq B_{\mathcal{C}} = \bigcup_{X \in \mathcal{C}} X$ there exists a $C'' \in \mathcal{C}$ such that $b(n) \in C''$. Now as $\mathcal{C}$ is a chain we have the following possibilities
   a. $(\boldsymbol{C' \subseteq C''})$ then $A \setminus \{b(n)\} \subseteq C' \subseteq C''$ and $b(n) \in C'' \Rightarrow A = A \setminus \{b(n)\} \bigcup \{b(n)\} \subseteq C''$ so if we take $C = C''$ we have $A \subseteq C \in S_{\mathcal{C}}$
   b. $(\boldsymbol{C'' \subseteq C'})$ then $A \setminus \{b(n)\} \subseteq C'$ and $b(n) \in C'' \subseteq C'$ so $A = A \setminus \{b(n)\} \bigcup \{b(n)\} \subseteq C'$, take then $C = C'$ and then we have $A \subseteq C \in S_{\mathcal{C}}$

   so in both cases we have found a $C \in \mathcal{C}$ such that $A \subseteq C$ proving that $n + 1 \in S_C$.

Using mathematical induction we have then that $S_{\mathcal{C}} = \mathbb{N}_0$ proving that 10.19 is indeed true ☐

Now if $\{\alpha_v\}_{v\in B_{\mathcal{C}}}$ is a family in F with finite support such that $0=\sum_{v\in B_{\mathcal{C}}}\alpha_v\cdot v$ then $\operatorname{support}(\{\alpha_v\}_{v\in B_{\mathcal{C}}})\subseteq B_{\mathcal{C}}$ is finite and we can use 10.19 to prove that:

$$\exists Y\in\mathcal{C}\text{ such that support}(\{\alpha_v\}_{v\in B_{\mathcal{C}}})\subseteq Y \tag{10.20}$$

As $Y\in\mathcal{C}\subseteq\mathcal{A}(R,W)$ we have by 10.15 that Y is linear independent. Define now $\{\beta_v\}_{v\in Y}$ by $\beta_v=\left\{\begin{smallmatrix}\alpha_v\text{ if }v\in B_{\mathcal{C}}\\ 0\text{ if }v\in Y\setminus B_{\mathcal{C}}\end{smallmatrix}\right.$ then we have trivially that $\operatorname{support}(\{\alpha_v\}_{v\in B_{\mathcal{C}}})=\operatorname{support}(\{\beta_v\}_{v\in Y})$ so that $\{\beta_v\}_{v\in Y}$ has finite support. Also $\sum_{v\in Y}\beta_v\cdot v=\sum_{v\in Y\setminus B_{\mathcal{C}}}\beta_v\cdot v+\sum_{v\in B_{\mathcal{C}}}\beta_v\cdot v=\sum_{v\in Y\setminus B_{\mathcal{S}}}0\cdot v+\sum_{v\in B_{\mathcal{C}}}\alpha_v\cdot v=0$ so by linear independence of Y that $\forall v\in Y$ we have $\beta_v=0$ and as $B_{\mathcal{C}}\subseteq Y$ we have $\forall v\in B_{\mathcal{C}}$ we have $\alpha_v=\beta_v=0$ proving that

$$B_{\mathcal{C}}\text{ is linear independent} \tag{10.21}$$

Using 10.21, 10.18 and the definition of $\mathcal{A}(R,W)$ (see 10.15) we have

$$B_{\mathcal{C}}\in\mathcal{A}(R,W) \tag{10.22}$$

Using 10.22 together with 10.17 let's us finally conclude that

$$\text{If }\mathcal{C}\text{ is }a\text{ non empty chain in }\mathcal{A}(R,W)\text{ then it is bounded above by }B_{\mathcal{C}} \tag{10.23}$$

If $\mathcal{C}=\emptyset$ is the empty chain then as $R\in\mathcal{A}(R,W)$ we have vacuously that $\forall C\in\mathcal{C}$ we have $C\subseteq R$ where $R\in\mathcal{A}(R,W)$ so R is a upper bound of $\mathcal{C}$. This together with 10.23 gives us

$$\text{Every chain in }\mathcal{A}(R,W)\text{ has }a\text{ upper bound in }\mathcal{A}(R,W) \tag{10.24}$$

Using Zorn's lemma (see 2.216) we have that

$$\exists M\in\mathcal{A}(R,W)\text{ such that }\forall X\in\mathcal{A}(R,W)\text{ we have }X\subseteq M \tag{10.25}$$

and by 10.15 we have also

$$R\subseteq M\subseteq W\text{ and }M\text{ is linear independent} \tag{10.26}$$

Let now $w\in W$ then we have the following cases to consider:

1. $\boldsymbol{(w\in M)}$ then as $M\subseteq S(M)$ (see 10.133) we have $w\in\mathcal{S}(M)$.
2. $\boldsymbol{(w\notin M)}$ then $\{w\}\bigcup M$ must be linear dependent [for if it is linear dependent then as $R\subseteq M\subseteq M\bigcup\{w\}\subseteq W$ we would have $M\bigcup\{w\}$ contradicting the maximality of M]. Hence there exists a $\{\alpha_v\}_{v\in M\bigcup\{w\}}$ in F with finite support such that $\sum_{v\in M\bigcup\{w\}}\alpha_v\cdot v=0$ and there exists a $v_0\in M\bigcup\{w\}$ such that $\alpha_{v_0}\neq 0$. Now $\operatorname{support}(\{\alpha_v\}_{v\in M\bigcup\{w\}})=\{v\in M\bigcup\{w\}|\alpha_v\neq 0\}=\{v\in M|\alpha_v\neq 0\}\bigcup\{v\in\{w\}|\alpha_v\neq 0\}=\operatorname{support}(\{a_v\}_{v\in M})\bigcup\{v\in\{w\}|\alpha_v\neq 0\}$. If now $\alpha_w=0$ then $\operatorname{support}(\{\alpha_v\}_{v\in M\bigcup\{w\}})=\operatorname{support}(\{\alpha_v\}_{v\in M})$ and as $0=\sum_{v\in M\bigcup\{w\}}\alpha_v\cdot w=\sum_{v\in M}\alpha_v\cdot v+\sum_{v\in\{w\}}\alpha_v\cdot v=\sum_{v\in M}\alpha_v\cdot v+0\cdot w=\sum_{v\in M}\alpha_v\cdot v$ and from independence from Mwe would then have that $\alpha_{v_0}=0$ a contradiction. So we must conclude that $\alpha_w\neq 0$ and then from $0=\sum_{v\in M\bigcup\{w\}}\alpha_v\cdot v=\sum_{v\in M}\alpha_v\cdot v+\sum_{v\in M}\alpha_v\cdot v+\sum_{v\in\{w\}}\alpha_v\cdot v=\sum_{v\in M}\alpha_v\cdot v+\alpha_w\cdot w\Rightarrow w=(-\alpha_w^{-1})\cdot\sum_{v\in M}\alpha_v\cdot v=\sum_{v\in M}(-\alpha_w^{-1}\cdot\alpha_v)\cdot v$ proving again that $w\in\mathcal{S}(M)$

From the above we conclude that $W \subseteq \mathcal{S}(M)$. Using 10.147 we have $V = \mathcal{S}(W) \subseteq \mathcal{S}(\mathcal{S}(M)) \underset{10.135}{=} \mathcal{S}(M) \subseteq V \Rightarrow \mathcal{S}(M) = V$ which together with 10.26 means that $\{m\}_{m \in M}$ is basis for V and $R \subseteq M \subseteq W$. □

Corollary 10.155. *Let V be a non trivial vector space over a field F and let $W \subseteq V$ be such that $\mathcal{S}(W) = V$ then there exist a set B with $B \subseteq W$ such that $\{b\}_{b \in B}$ forms a basis of V. As $\mathcal{S}(V) = V$ we see directly that every non trivial vector space has a basis.*

Proof. As $V \neq \{0\}$ there exists a $v_0 \in V$ such that $v_0 \neq 0$, as $V = \mathcal{S}(W)$ there exist a $\{\alpha_v\}_{v \in W}$ a family in F with finite support so that $v_0 = \sum_{w \in W} \alpha_w \cdot w$. Now if $W = \{0\}$ or $W = \emptyset$ then $\mathrm{support}(\{\alpha_w \cdot w\}_{w \in W}) = \emptyset$ and thus $v_0 = \sum_{w \in W} \alpha_w \cdot w = 0$ a contradiction so there must exists a $w \in W$ with $w \neq 0$. If now there exists a $\{\alpha_v\}_{v \in \{w\}}$ such that $0 = \sum_{v \in \{w\}} \alpha_v \cdot v = \alpha_w \cdot w \Rightarrow \alpha_w = 0 \Rightarrow \forall v \in \{w\}$ we have $\alpha_w = 0$ proving that $\{w\} \subseteq W$ is a linear independent set. Using the previous theorem we have then to find a linear independent set B with $\{w\} \subseteq B \subseteq W$ such that $\{b\}_{b \in B}$ is a basis of V. □

Corollary 10.156. *Let V be a non trivial vector space over a field F and let $\{v_i\}_{i \in I}$ be a family in V such that $\mathcal{S}(\{v_i | i \in I\}) = V$ then there exists a $J \subseteq I$ such that $\{v_i\}_{i \in J}$ is a basis of V [equivalent there exists a $\{w_i\}_{i \in J}$ such that $\{w_i\}_{i \in J}$ is a basis and $\{w_i\}_{i \in J} \preccurlyeq \{v_i\}_{i \in I}$].*

Proof. First as $v = \{v_i\}_{i \in I}$ is a family of elements in V we have by definition of a family of elements (see 2.100) that it is a function $v: I \rightarrow V$ with $i \rightarrow v(i) = v_i$. We can use 2.225 to find a subset $J' \subseteq I$ so that $v_{|J'}: J' \rightarrow v(I) = \{v_i | i \in I\}$ is a bijection. Now as $\mathcal{S}(\{v_i | i \in I\}) = V$ we can use 10.155 to find a set B with $B \subseteq \{v_i | i \in I\}$ such that $\{b\}_{b \in B}$ forms a basis for V. Take now $J = (v_{|J'})^{-1}(B) \subseteq J'$ then $(v_{|J'})_{|J} = v_{|J}: J \rightarrow B$ forms a bijection such that $\forall j \in J$ we have $v_{|J}(j) = v(j) = v_j$ and defines the family $\{v_i\}_{i \in J} \preccurlyeq \{v_i\}_{i \in I}$. We prove now that $\{v_i\}_{i \in J}$ forms a basis. Now if $\{\alpha_i\}_{i \in J}$ is a family in F with finite support such that $0 = \sum_{i \in J} \alpha_i \cdot v_i$ define then $\{\beta_b\}_{b \in B}$ such that $\beta_b = \alpha_{(v_{|J})^{-1}(b)}$ then $\beta_{v_i} = \beta_{v_{|J}(i)} = \alpha_{(v_{|J})^{-1}(v_J(i))} = \alpha_i$ and thus $0 = \sum_{i \in J} \beta_{v_i} \cdot v_i \underset{v_{|J}: J \rightarrow B \text{ is bijective and } 10.44}{=} \sum_{b \in B} \beta_b \cdot b$ and as $\{b\}_{b \in B}$ is linear independent we have $\forall b \in B$ that $\beta_b = 0$, now if $i \in J$ then $\alpha_i = \beta_{v_i} = 0$ so $\forall i \in J$ we have $\alpha_i = 0$ proving that $\{v_i\}_{i \in J}$ is linear independent. As also $\{v_i | i \in J\} = v_{|J}(J) = B$ we have $\mathcal{S}(\{v_i | i \in J\}) = \mathcal{S}(B) = V$ and all conditions are fulfilled to make $\{v_i\}_{i \in J}$ a basis of V. □

Corollary 10.157. *Let V be a non trivial vector space over a field F then if $W \subseteq V$ is a linear independent subset of V then there exists a set B with $W \subseteq B$ and for which $\{b\}_{b \in B}$ is a basis of V.*

Proof. As $\mathcal{S}(V) = V$ and $W \subseteq V$ we can use 10.154 to find a set $W \subseteq B \subseteq F$ such that $\{b\}_{b \in B}$ is a basis for V. □

The following is a useful theorem if we have a finite dimensional basis to express it as a family based on the indexing set $\{1, \ldots, n\}$

Theorem 10.158. *If V is a vector space over a field F and $B \subseteq V$ a set such that $\{b\}_{b\in B}$ is a basis of V and $\#(B)=n\in\mathbb{N}$ (so B is finite) then there exists a basis $\{v_i\}_{i\in\{1,\ldots,n\}}$ in V*

Proof. If $\#(B)=n>0$ then there exists a bijection $v:\{1,\ldots,n\}\rightarrow B$ which defines the family $\{v_i\}_{i\in\{1,\ldots,n\}}$. If now $\{\alpha_i\}_{i\in\{1,\ldots,n\}}$ is a family in F with finite support such that $\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot v_i=0$, define then $\{\beta_b\}_{b\in B}$ by $\beta_b=\alpha_{v^{-1}(b)}$ then we have $\beta_{v_i}=\beta_{v(i)}=\alpha_{v^{-1}(v(i))}=\alpha_i$ and thus $0=\sum_{i\in\{1,\ldots,n\}}\beta_{v_i}\cdot v_i \underset{10.44\text{ and }v\text{ is a bijection}}{=} \sum_{b\in B}\beta_b\cdot b$. By linear independence of $\{b\}_{b\in B}$ we have $\forall b\in B$ that $\beta_b=0$. If now $i\in\{1,\ldots,n\}$ then $\alpha_i=\beta_{v(i)}=0$ so we have proved that $\{v_i\}_{i\in I}$ is linear independent. Finally $\mathcal{S}(\{v_i|i\in\{1,\ldots,n\}\})=\mathcal{S}(B)=V$ so $\{v_i\}_{i\in I}$ spans V. □

Definition 10.159. (Kronicker Delta) *Let $n\in\mathbb{N}$ and F a field with unit element 1 and neutral element 0 then we define $\delta_n:\{1,\ldots,n\}\times\{1,\ldots,n\}\rightarrow F$ by $(i,j)\rightarrow\delta_n(i,j)=\left\{\begin{smallmatrix}1\text{ if }i=j\\0\text{ if }i\neq j\end{smallmatrix}\right.$. If n is understood to be know then we write just $\delta_{i,j}$ instead of $\delta_n(i,j)$.*

Theorem 10.160. *If $n\in N_0$, F is field and $\{\alpha_i\}_{i\in\{1,\ldots,n\}}$ a family in F then $\sum_{i=1}^n\delta_{i,j}\cdot\alpha_i=\alpha_j$*

Proof. $\sum_{i=1}^n\delta_{i,j}\cdot\alpha_i=\sum_{i\in\{1,\ldots,n\}}\delta_{i,j}\cdot\alpha_i=\sum_{i\in\{1,\ldots,n\}\backslash\{j\}}\delta_{i,j}\cdot\alpha_i+\sum_{i\in\{j\}}\delta_{i,j}\cdot\alpha_i=\sum_{i\in\{1,\ldots,n\}\backslash\{j\}}0\cdot\alpha_i+\delta_{j,j}\cdot\alpha_j=0+1\cdot\alpha_j=\alpha_j$ □

Definition 10.161. *If F is a field, $n\in\mathbb{N}$ and F^n the vector space defined over F (see 10.116 for the definition) define then the family $\{\mathcal{E}_i\}_{i\in\{1,\ldots,n\}}$ in F^n by $(\mathcal{E}_i)_j=\mathcal{E}_i(j)=\delta_{i,j}$ $\forall i,j\in\{1,\ldots,n\}$*

Theorem 10.162. *If F is a field $n\in\mathbb{N}$ and F^n the vector space over F defined in 10.116 then $\{\mathcal{E}_i\}_{i\in\{1,\ldots,n\}}$ is a basis of F^n*

Proof. If $x=(x_1,\ldots,x_n)\in F^n$ then $x=\{x_i\}_{i\in\{1,\ldots,n\}}$ is a family in F and $\forall i\in\{1,\ldots,n\}$ we have $\left(\sum_{j\in\{1,\ldots,n\}}x_j\cdot\mathcal{E}_j\right)_i=(\sum_{j=1}^n x_j\cdot\mathcal{E}_j)_i \underset{10.10}{=} \sum_{j=1}^n(x_j\cdot\mathcal{E}_j)_i=\sum_{j=1}^n x_j\cdot(\mathcal{E}_j)_i=\sum_{j=1}^n x_j\cdot\delta_{j,i} \underset{10.160}{=} x_i$ so we have $x=\sum_{j\in\{1,\ldots,n\}}x_j\cdot\mathcal{E}_j$ and thus we have $\mathcal{S}(\{\mathcal{E}_i|i\in\{1,\ldots,n\}\})=F^n$.

Now to prove linear independence assume that there exists a $\{\alpha_j\}_{j\in\{1,\ldots,n\}}$ such that $0=(0,\ldots,0)=\sum_{j\in\{1,\ldots,n\}}\alpha_j\cdot\mathcal{E}_j$ then we have $0=\left(\sum_{j\in\{1,\ldots,n\}}\alpha_j\cdot\mathcal{E}_j\right)_i=(\sum_{j=1}^n\alpha_j\cdot\mathcal{E}_j)_i \underset{10.10}{=} \sum_{j=1}^n(\alpha_j\cdot\mathcal{E}_j)_i=\sum_{j=1}^n\alpha_j\cdot(\mathcal{E}_j)_i=\sum_{j=1}^n\alpha_j\cdot\delta_{\text{ji}}=\alpha_i$ proving that $\forall i\in\{1,\ldots,n\}$ we have $\alpha_i=0$ □

Applying the above to the fields $\mathbb{R}$ and $\mathbb{C}$ we have the following corollaries:

Corollary 10.163. *$\{\mathcal{E}_i\}_{i\in\{1,\ldots,n\}}$ is a basis of the vector space $\mathbb{R}^n$ over the field $\mathbb{R}$*

Corollary 10.164. *$\{\mathcal{E}_i\}_{i\in\{1,\ldots,n\}}$ is a basis of the vector space $\mathbb{C}^n$ over the field $\mathbb{C}$*

We proceede now to construct the basis for a product of finite vector spaces based of these vector spaces. First we need a little lemma.

Lemma 10.165. *Let $n\in\mathbb{N}$, $\{n_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{N}$ then if we define $\{N_i\}_{i\in\{1,\ldots,n\}}$ by $N_i=\{(i,j)|j\in(1,\ldots,n_i)\}$ we have*

1. *The function β: $\bigcup_{i\in\{1,\ldots,n\}} N_i\rightarrow\{1,\ldots,\sum_{i=1}^n n_i\}$ by $\beta(i,j)=(\sum_{k=1}^{i-1} n_k)+j$ is a bijection*
2. *The inverse bijection $\tau=\beta^{-1}$: $\{1,\ldots,\sum_{i=1}^n n_i\}\rightarrow\bigcup_{i\in\{1,\ldots,n\}} N$ satisfies $\forall k\in\{1,\ldots,\sum_{i=1}^n n_i\}$ that $\tau(k)_1\in\{1,\ldots,n\}$, $\tau(k)\in\{1,\ldots,n_{\tau(k)_1}\}$ and $\forall l\in\{1,\ldots,n\}$ we have if $i\in\tau^{-1}(N_l)$ that $\tau(i)_1=l$*
3. *$\forall k\in\{1,\ldots,n\}$ we have that the function γ_k: $\tau^{-1}(N_k)\rightarrow\{1,\ldots,n_k\}$ defined by $\gamma_k(i)=\tau(i)_2$ is a bijection.*

Proof.

1. If $(i,j)\in\bigcup_{k\in\{1,\ldots,n\}} N_k$ then $\exists k\in\{1,\ldots,n\}$ such that $(i,j)\in N_k$ hence $(i,j)\in\{(k,l)|l\in\{1,\ldots,n_k\}\}$ proving that $i=k$ and $j\in\{1,\ldots,n_k\}=\{1,\ldots,n_i\}$. This gives
$$\forall(i,j)\in\bigcup_{k\in\{1,\ldots,n\}} N_k\Rightarrow j\in\{1,\ldots,n_i\}\text{ and }i\in\{1,\ldots,n\} \tag{10.27}$$
injectivity. Let $(i,j),(r,s)\in\bigcup_{k\in\{1,\ldots,n\}} N_i$ be such that $\beta(i,j)=\beta(r,s)$ then we have $(\sum_{k=1}^{i-1} n_k)+j=(\sum_{k=1}^{r-1} n_k)+s$. Assume that $i\neq r$ then we may assume without loosing generality that $i<r$ (otherwise exchange i with r). Using 10.27 we have that $j\in\{1,\ldots,n_i\}$ hence $(\sum_{k=1}^{i-1} n_k)+j\leqslant(\sum_{k=1}^{i-1} n_k)+n_i=\sum_{k=1}^{i} n_k\leqslant_{i<r\Rightarrow i\leqslant r-1}\sum_{k=1}^{r-1} n_k<(\sum_{k=1}^{r-1} n_k)+s$ proving that $(\sum_{k=1}^{i-1} n_k)+j<(\sum_{k=1}^{r-1} n_k)+s$ in contradiction with $(\sum_{k=1}^{i-1} n_k)+j=(\sum_{k=1}^{r-1} n_k)+s$. So we must have that $i=r$ and thus $(\sum_{k=1}^{i-1} n_k)+j=(\sum_{k=1}^{i-1} n_k)+s\Rightarrow j=1$, hence $(i,j)=(r,s)$ proving injectivity.

   surjectiviy. If $i\in\{1,\ldots,\sum_{k=1}^n n_k\}$ then as $\sum_{i=1}^{1-1} n_k=\sum_{i=1}^{0} n_k=0<i$ we have that $\mathcal{S}=\{k\in\{1,\ldots,n\}|\sum_{l=1}^{k-1} n_l<i\}\neq\emptyset$ hence $i_m=\max(\mathcal{S})$ exists. We have now two possible cases for i_m:

   $i_m=n$. hence $\sum_{l=1}^{n-1} n_l<i\leqslant\sum_{l=1}^{n} n_l$ then $0<i-\sum_{l=1}^{n-1} n_l\leqslant\sum_{l=1}^{n} n_l-\sum_{l=1}^{n-1} n_l=n_n$ so that if we take $s=i-\sum_{l=1}^{n-1} n_l$ we have that $s\in\{1,\ldots,n_n\}$ and $i=\sum_{l=1}^{n-1} n_l+s$ proving that $(i,s)\in N_n$ and $\beta(s)=i$

   $i_m<n$. hence $i_m+1\notin\mathcal{S}$ so that $\sum_{l=1}^{i_m-1} n_l<i\leqslant\sum_{l=1}^{(i_m+1)-1} n_l=\sum_{l=1}^{i_m} n_l$ then $0<i-\sum_{l=1}^{i_m-1} n_l\leqslant\sum_{l=1}^{i_m} n_l-\sum_{l=1}^{i_m-1} n_l=n_{i_m}$ so that if we take $s=i-\sum_{l=1}^{i_m-1} n_l$ we have that $s\in\{1,\ldots,n_{i_m}\}$ and $i=\sum_{l=1}^{i_m-1} n_l+s$ proving that $(i,s)\in N_{i_m}$ and $\beta(s)=i$

2. If $k\in\{1,\ldots,\sum_{i=1}^n N_i\}$ then $\tau(k)=(\tau(k)_1,\tau(k)_2)\in\bigcup_{k\in\{1,\ldots,n\}} N_k$ hence using 10.27 we have that $\tau(k)_1\in\{1,\ldots,n\}$ and $\tau(k)_2\in\{1,\ldots,n_{\tau(k)_1}\}$. If now $l\in\{1,\ldots,n\}$ and $i\in\tau^{-1}(N_l)$ then $(\tau(i)_1,\tau(i)_2)\in N_l=\{(l,i)|i\in\{1,\ldots,n_l\}\}$ so that $\tau(i)_1=l$.

3. Let $k \in \{1, \ldots, n\}$ then we have

   injectivity. If $i, j \in \tau^{-1}(N_k)$ and $\gamma_k(i) = \gamma_k(j)$ then $\tau_k(i)_2 = \tau_k(j)_1$ and as by (2) we have that $\tau(i)_1 = k = \tau(j)_1$ we conclude that $\tau(i) = \tau(j)$ hence $i = j$.

   surjectivity. Let $l \in \{1, \ldots, n_k\}$ then $(k, l) \in N_k$ so as τ is a bijection there exists a $i \in \{1, \ldots, \sum_{j=1}^{n} n_j\}$ such that $\tau(i) = (k, l) \in N_k$. By definition we have then that $i \in \tau^{-1}(N_k)$ and $\gamma_k(i) = \tau(i)_2 = l \in \{1, \ldots, n_k\}$ □

Theorem 10.166. *Let $n \in \mathbb{N}$, $\{V_i\}_{i \in \{1,\ldots,n\}}$ be a finite family of non trivial vector spaces over a field F with $\forall i \in \{1, \ldots, n\}$ $\{e_{i,j}\}_{j \in \{1,\ldots,n_i\}}$ a basis for V_i. Given $k \in \{1, \ldots, \sum_{i=1}^{n} n_i\}$ define $f_k = ((f_k)_1, \ldots, (f_k)_n) \in \prod_{i \in \{1,\ldots,n\}} V_i$ by $(f_k)_i = e_{\tau(k)_1, \tau(k)_2}$ if $\tau(k)_1 = i$ and $(f_k)_i = 0$ if $\tau(k)_1 \in \{1, \ldots, n\} \setminus \{i\}$ (here $\tau: \{1, \ldots, \sum_{i=1}^{n} n_i\} \rightarrow \bigcup_{i \in \{1,\ldots,n\}} N_i$ is defined in the previous lemma, which also ensures that $\tau(k)_1 \in \{1, \ldots, n\}$ and $\tau(k)_2 \in \{1, \ldots, n_{\tau(k)_1}\}$ as it should). Then $\{f_k\}_{k \in \{1,\ldots,\sum_{i=1}^{n} n_i\}}$ is a basis of $\prod_{i \in \{1,\ldots,n\}} V_i$ (which is a vector space by 10.118)*

Proof. Define $N = \sum_{i=1}^{n} n_i$. Then if $k \in \{1, \ldots, n\}$ and $l \in \tau^{-1}(N_k)$ then $\tau(l) \in N_k$ and thus $\tau(l)_1 = k$ and $\tau(l)_2 \in \{1, \ldots, n_k\}$ so that $(f_l)_k = e_{k, \tau(l)_2} \in \{e_{k,j} | j \in \{1, \ldots, n_k\}\}$. If $l \in \{1, \ldots, N\} \setminus \tau^{-1}(N_k)$ then $\tau(l)_1 \neq k$ [if $\tau(l)_1 = k$ then by the previous lemma we have that $\tau(l)_2 \in \{1, \ldots, n_k\}$ hence $\tau(l) \in N_k \Rightarrow l \in \tau^{-1}(k)$ a contradiction], so $(f_l)_k = 0$. To summarize we have proved that

$$\forall k \in \{1, \ldots, n\} \vDash \text{if } l \in \tau^{-1}(N_k) \text{ then } (f_l)_k = e_{k, \tau(l)_2} \in \{e_{k,j} | j \in \{1, \ldots, n_k\}\} \text{ and if } l \in \{1, \ldots, N\} \setminus \tau^{-1}(N_k) \text{ then } (f_l)_k = 0 \tag{10.28}$$

Let now $\{\alpha\}_{i \in \{1,\ldots,N\}} \subseteq F$ then we have given $k \in \{1, \ldots, n\}$ that

$$\begin{aligned}
\left(\sum_{i=1}^{N} \alpha_i \cdot f_i\right)_k &= \sum_{i=1}^{N} \alpha_i \cdot (f_i)_k \\
&= \sum_{i \in \{1,\ldots,N\}} \alpha_i \cdot (f_i)_k \\
&= \sum_{i \in \{1,\ldots,N\} \setminus \tau^{-1}(N_k)} \alpha_i \cdot (f_i)_k + \sum_{i \in \tau^{-1}(N_k)} \alpha_i \cdot (f_i)_k \\
&\underset{10.28}{=} 0 + \sum_{i \in \tau^{-1}(N_k)} \alpha_i \cdot e_{k, \tau(i)_2} \\
&= \sum_{i \in \tau^{-1}(N_k)} \alpha_i \cdot e_{k, \tau(i)_2}
\end{aligned}$$

proving that

$$\forall k \in \{1, \ldots, n\} \vDash \left(\sum_{i=1}^{N} \alpha_i \cdot f_i\right)_k = \sum_{i \in \tau^{-1}(N_k)} \alpha_i \cdot e_{k, \tau(i)_2} \tag{10.29}$$

If now $\{\alpha_i\}_{i\in\{1,\ldots,N\}}\subseteq F$ is such that $0=\sum_{i=1}^N \alpha_i\cdot f_i$ we have by the above that $\forall k\in\{1,\ldots,n\}$ that $0=\sum_{i\in\tau^{-1}(N_k)}\alpha_i\cdot e_{k,\tau(i)_2}$. As $\{e_{k,j}\}_{j\in\{1,\ldots,n_k\}}$ is a basis we have using 10.28 that $\forall i\in\tau^{-1}(N_k)\quad \alpha_i=0$. If $m\in\{1,\ldots,N\}$ we have that $\tau(m)\in\bigcup_{k\in\{1,\ldots,n\}}N_k$ so that there exists a $k\in\{1,\ldots,n\}$ such that $\tau(m)\in N_k\Rightarrow m\in\tau^{-1}(N_k)$ hence $\alpha_m=0$. So we proved that $\forall m\in\{1,\ldots,N\}$ we have $\alpha_m=0$ and thus that

$$\{f_i\}_{i\in\{1,\ldots,N\}}\text{ is linear independent} \tag{10.30}$$

To prove completencess let $x=(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}V_i$ then we have $\forall k\in\{1,\ldots,n\}$ that there exists a $\{\alpha_{k,i}\}_{j\in\{1,\ldots,n_k\}}$ such that $x_k=\sum_{i=1}^{n_k}\alpha_{k,i}\cdot e_{k,i}$. Define now $\{\beta_k\}_{k\in\{1,\ldots,N\}}$ by $\beta_k=\alpha_{\tau(k)_1,\tau(k)_2}$. We have then $\forall k\in\{1,\ldots,n\}$ that if $i\in\tau^{-1}(N_k)$ that $(\tau(i)_1,\tau(i)_2)=\tau(i)\in N_k\Rightarrow\tau(i)_1=k\wedge\tau(i)_2\in\{1,\ldots,n_k\}$ so that $\beta_i=\alpha_{\tau(i)_1,\tau(i)_2}=\alpha_{k,\tau(i)_2}$ giving

$$\forall k\in\{1,\ldots,n\}\text{ we have if } i\in\tau^{-1}(N_k)\text{ that }\beta_k=\alpha_{k,\tau(i)_2} \tag{10.31}$$

hence we have $\forall k\in\{1,\ldots,n\}$ that

$$\begin{aligned}\left(\sum_{i=1}^N\beta_i\cdot f_i\right)_k &\underset{10.29}{=} \sum_{i\in\tau^{-1}(N_k)}\beta_i\cdot e_{k,\tau(i)_2}\\ &\underset{10.31}{=} \sum_{i\in\tau^{-1}(N_k)}\alpha_{k,\tau(i)_2}\cdot e_{k,\tau(i)_2}\end{aligned}$$

If we use now the bijection $\gamma_k:\tau^{-1}(N_k)\rightarrow\{1,\ldots,n_k\}$ defined by $\gamma_k(i)=\tau(i)_2$ (see the previous lemma 10.165 (3)) then we have that $(\sum_{i=1}^N\beta_i\cdot f_i)_k=\sum_{i\in\tau^{-1}(N_k)}\alpha_{k,\gamma_k(i)}\cdot e_{k,\gamma_k(i)}\underset{10.44}{=}\sum_{i\in\{1,\ldots,n_k\}}\alpha_{k,i}\cdot e_{k,i}=x_k$. So we have proved that $\forall\kappa\{1,\ldots,n\}$ we have that $(\sum_{i=1}^N\beta_i\cdot f_i)_k=x_k$ where $x=(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots n\}}V_i$ hence $x=\sum_{i=1}^N\beta_i\cdot f_i$ proving that $\{f_i\}_{i\in\{1,\ldots,N\}}$ is covering $\prod_{i\in\{1,\ldots,n\}}V_i$. □

10.5 Free vector space over a set

In this section we show that given a non empty set we can construct a vector space that has this set as a basis.

Definition 10.167. *Given a non empty set X and a field $\langle F,+,\cdot\rangle$ define then $\mathcal{F}(X,F)=\{f:X\rightarrow\mathbb{K}|f^{-1}(F\setminus\{0\})\text{ is finite}\}$. In other words $\mathcal{F}(X,F)$ is the set of functions from X to F that are nonzero only for a finite number of points in X.*

Definition 10.168. *Given a set X and a field F then we define*

1. *If $f,g\in\mathcal{F}(X,F)$ then $f+g:X\rightarrow F$ is defined by $x\rightarrow f(x)+g(x)$*
2. *If $\alpha\in F$ and $f\in\mathcal{F}(X,F)$ then $\alpha\cdot f:X\rightarrow F$ is defined by $x\rightarrow\alpha\cdot f(x)$*
3. *$\forall a\in X$ we define $\delta_a\in\mathcal{F}(X,F)$ by $\delta_a(x)=\left\{\begin{array}{l}1\text{ if }x=a\\0\text{ if }x\in X\setminus\{a\}\end{array}\right.$*

Theorem 10.169. *Given a non empty set X and a field $\langle F,+,\cdot\rangle$ then $\forall f,g\in\mathcal{F}(X,F)$ and $\alpha\in F$ we have that $\langle\mathcal{F}(X,F),+,\cdot\rangle$ is a vector space over F (see previous definition for $+,\cdot$) with basis $\{\delta_a\}_{a\in X}$ and the zero element $0\colon X\rightarrow F$ defined by $x\rightarrow 0$. This vector space is called the* ***free vector space over a set****. The set X is embedded in $\mathcal{F}(X,F)$ by the injection $\delta\colon X\rightarrow\mathcal{F}(X,F)$ defined by $a\rightarrow\delta_a$ if $\langle F,+,\cdot\rangle$ has characterization 0.*

Proof. If $f,g\in\mathcal{F}(X,F)$ then if $x\in(f+g)^{-1}(F\setminus\{0\})$ we have $(f+g)(x)\neq 0\Rightarrow f(x)+g(x)\neq 0\Rightarrow f(x)\neq 0\vee g(x)\neq 0\Rightarrow x\in f^{-1}(F\setminus\{0\})\bigcup g^{-1}(F\setminus\{0\})\Rightarrow(f+g)^{-1}(F\setminus\{0\})\subseteq f^{-1}(F\setminus\{0\})\bigcup g^{-1}(F\setminus\{0\})$ which means as $f^{-1}(F\setminus\{0\})$, $g^{-1}(F\setminus\{0\})$ are finite that $(f+g)^{-1}(F\setminus\{0\})$ is finite and thus that $f+g\in\mathcal{F}(X,F)$ proving that $+\colon\mathcal{F}(X,F)\times\mathcal{F}(X,F)\rightarrow\mathcal{F}(X,F)$ is well defined.

Also if $\alpha\in F$ and $f\in\mathcal{F}(X,F)$ then if $x\in(\alpha\cdot f)^{-1}(F\setminus\{0\})$ we have $\alpha\cdot f(x)\neq 0\Rightarrow f(x)\neq 0\Rightarrow x\in f^{-1}(F\setminus\{0\})\Rightarrow(\alpha\cdot f)^{-1}(F\setminus\{0\})\subseteq f^{-1}(F\setminus\{0\})$ which as $f^{-1}(F\setminus\{0\})$ is finite proves that $\alpha\cdot f\in\mathcal{F}(X,F)$ and thus that $\cdot\colon F\times\mathcal{F}(X,F)\rightarrow\mathcal{F}(X,F)$ is well defined. Now to prove that it is a vector space:

1. $\langle\mathcal{F}(X,F),+\rangle$ is a Abelian group
   a. **(Associativity)** If $f,g,h\in\mathcal{F}(X,F)$ then $\forall x\in X$ we have $(f+(g+h))(x)=f(x)+(g+h)(x)=f(x)+(g(x)+h(x))=(f(x)+g(x))+h(x)=(f+g)(x)+h(x)=((f+g)+h)(x)$ proving that $f+(g+h)=(f+g)+h$
   b. **(Commutativity)** If $f,g\in\mathcal{F}(X,F)$ then $\forall x\in X$ we have $(f+g)(x)=f(x)+g(x)=g(x)+f(x)=(g+f)(x)$ proving that $f+g=g+f$
   c. **(Neutral Element)** If $f\in\mathcal{F}(X,f)$ then $(f+0)(x)=f(x)+0(x)=f(x)$ giving $f+0=f\underset{\text{commutativity}}{=}0+f$
   d. **(Inverse Element)** If $f\in\mathcal{F}(X,F)$ define then $-f\colon X\rightarrow F$ by $x\rightarrow -f(x)$ which as $x\in(-f)^{-1}(F\setminus\{0\})$ then $(-f)(x)\neq 0\Rightarrow -f(x)\neq 0\Rightarrow f(x)\neq 0\Rightarrow x\in f^{-1}(F\setminus\{0\})\Rightarrow(-f)^{-1}(F\setminus\{0\})\subseteq f^{-1}(F\setminus\{0\})$ which is finite proving that $-f\in\mathcal{F}(X,F)$. Now $\forall x\in X$ we have $(f+(-f))(x)=f(x)+(-(f(x)))=0=0(x)$ proving that $f+(-f)=0\underset{\text{commutativity}}{=}(-f)+f=0$
2. Vector spaces axioms
   a. If $\alpha\in F$ and $f,g\in\mathcal{F}(X,F)$ then $\forall x\in X$ we have $(\alpha\cdot(f+g))(x)=\alpha\cdot(f+g)(x)=\alpha\cdot(f(x)+g(x))=\alpha\cdot f(x)+\alpha\cdot g(x)=(\alpha\cdot f)(x)+(\beta\cdot g)(x)=((\alpha\cdot f)+(\beta\cdot g))(x)$ proving that $\alpha\cdot(f+g)=\alpha\cdot f+\alpha\cdot g$
   b. If $\alpha,\beta\in F$ and $f\in\mathcal{F}(X,F)$ then $\forall x\in X$ we have $((\alpha+\beta)\cdot f)(x)=(\alpha+\beta)\cdot f(x)=\alpha\cdot f(x)+\beta\cdot f(x)=((\alpha\cdot f)(x)+(\beta\cdot f)(x))=((\alpha\cdot f)+(\beta\cdot f))(x)$ so that $(\alpha+\beta)\cdot f=\alpha\cdot f+\beta\cdot f$
   c. If $\alpha,\beta\in F$ and $f\in\mathcal{F}(X,F)$ then $\forall x\in X$ we have $((\alpha\cdot\beta)\cdot f)(x)=(\alpha\cdot\beta)\cdot f(x)=\alpha\cdot(\beta\cdot f(x))=\alpha\cdot((\beta\cdot f)(x))=(\alpha\cdot(\beta\cdot f))(x)$ giving that $(\alpha\cdot\beta)\cdot f=\alpha\cdot(\beta\cdot f)$
   d. If $f\in\mathcal{F}(X,F)$ then $\forall x\in X$ we have $(1\cdot f)(x)=1\cdot f(x)=f(x)$ proving that $1\cdot f=f$

Next we have to prove that $\{\delta_a\}_{a\in X}$ is a basis of $\mathcal{F}(X,F)$

1. If $f\in\mathcal{F}(X,F)$ then $f^{-1}(F\setminus\{0\})$ is finite and we have two cases to consider
   a. $\boldsymbol{(f^{-1}(F\setminus\{0\})=\emptyset)}$ then $f=0$ and as X is non empty there exists a $a\in X$ and we have $f=0\cdot\delta_a=0$
   b. $\boldsymbol{(f^{-1}(F\setminus\{0\})\neq\emptyset)}$ then there exists a bijection $b\colon\{1,\ldots,n\}\rightarrow f^{-1}(F\setminus\{0\})$ defining a family $\{b_i\}_{i\in\{1,\ldots,n\}}$ such that $\forall y\in f^{-1}(F\setminus\{0\})$ we have that $\exists! i\in\{1,\ldots,n\}$ such that $b_i=y$ and $\forall i,j\in\{1,\ldots,n\}$ we have $i\neq j\Rightarrow b_i\neq b_j$ Take now $\sum_{i\in\{1,\ldots,n\}} f(b_i)\cdot\delta_{b_i}$ then we have for $x\in X$
      i. $\boldsymbol{(f(x)=0)}$ then $x\notin f^{-1}(F\setminus\{0\})$ so $\forall i\in\{1,\ldots,n\}$ we have $b_i\neq x$ and thus $\delta_{b_i}(x)=0$ giving that $\left(\sum_{i\in\{1,\ldots,n\}} f(b_i)\cdot\delta_{b_i}\right)(x)=\sum_{i\in\{1,\ldots,n\}} f(b_i)\cdot\delta_{b_i}(x)=\sum_{i\in\{1,\ldots,n\}} f(b_i)\cdot 0=0=f(x)$.
      ii. $\boldsymbol{(f(x)\neq 0)}$ then $x\in f^{-1}(F\setminus\{0\})$ so there exists a unique $i\in\{1,\ldots,n\}$ such that $b_i=x$ and thus for $j\in\{1,\ldots,n\}$ we have $\delta_j(x)=\delta_j(b_i)=\delta_{i,j}$ so that $\left(\sum_{j\in\{1,\ldots,n\}} f(b_j)\right)(x)=\sum_{j\in\{1,\ldots,n\}} f(b_j)\cdot\delta_{b_j}(x)=\sum_{j\in\{1,\ldots,n\}} f(b_j)\cdot\delta_{i,j}=f(b_i)=f(x)$

      proving that $\sum_{i\in\{1,\ldots,n\}} f(b_i)\cdot\delta_{b_i}=f$

   So $\mathcal{S}(\{\delta_a\}_{a\in X})=\mathcal{F}(X,F)$
2. To prove linear independence, let $\{\alpha_a\}_{a\in X}$ be a family in F with finite support such that $\sum_{a\in X}\alpha_a\cdot\delta_a=0$ then $\forall x\in X$ we have $0=\sum_{a\in X}\alpha_a\cdot\delta_a(x)\underset{\delta_a(x)=0\,\text{if}\,x\neq 0}{=}\alpha_x\cdot\delta_x(x)=\alpha_x\Rightarrow\alpha_x=0\Rightarrow\forall x\in X$ we have $\alpha_x=0$.

Finally we have to prove that $\delta\colon X\rightarrow\mathcal{F}(X,F)$ defined by $a\rightarrow\delta_a$ is a injection, so assume that $\delta_{a_1}=\delta_{a_2}$ then $\forall x\in X$ we have $\delta_{a_1}(x)=\delta_{a_2}(x)$ so if $a_1\neq a_2$ then $1=\delta_{a_1}(a_1)=\delta_{a_2}(a_1)=0$ which is a contradiction if F has characterization 0 so we must have that $a_1=a_2$. □

10.6 Linear Mappings

Definition 10.170. *If X and Y are vector spaces over a field F then a function L: $X\rightarrow Y$ is a **linear mapping** if $\forall\alpha,\beta\in F$ and $\forall x,y\in X$ we have $L(\alpha\cdot x+\beta\cdot y)=\alpha\cdot L(x)+\beta\cdot L(y)$*

Example 10.171. Let $n\in\mathbb{N}$, $\{X_i\}_{i\in\{1,\ldots,n\}}$ a family of vector spaces over a field F then given $i\in\{1,\ldots,n\}$ we have that $\pi_i\colon\prod_{j\in\{1,\ldots,n\}} X_j\rightarrow X_i$ is linear

Proof. This follows trivially of the definition of π_i □

Definition 10.172. *If X and Y are vector spaces over a field F then a linear mapping $L\colon X\rightarrow Y$ is regular if L is injective. and the vector space $\prod_{i\in\{1,\ldots,n\}} X_i$ (see 10.118)*

Lemma 10.173. *Let X and Y be vector spaces over a field F and $L\colon X\rightarrow Y$ is a linear mapping then $L(0)=0$*

Proof. $L(0)=L(0\cdot 0)=0\cdot L(0)=0$ □

Theorem 10.174. *Let X and Y be vector spaces over a field F and $L: X\to Y$ a linear mapping then*

1. *If $n\in\mathbb{N}$ and $\{\alpha_i\}_{i\in\{1,\ldots,n\}}$, $\{x_i\}_{i\in\{1,\ldots,n\}}$ are families in F and X then $L(\sum_{i=1}^n \alpha_i\cdot x_i)=\sum_{i=1}^n \alpha_i\cdot L(x_i)$*
2. *If $\{\alpha_i\}_{i\in I}$ is a family in F with finite support and $\{x_i\}_{i\in I}$ a family in X then $L(\sum_{i\in I}\alpha_i\cdot x_i)=\sum_{i\in I}\alpha_i\cdot L(x_i)$*
3. *If $A\subseteq X$ and $\{\alpha_a\}_{a\in A}$ then $L(\sum_{a\in A}\alpha_a\cdot a)=\sum_{a\in A}\alpha_a\cdot L(a)$*

Proof.

1. *This is proved by induction, let $S=\{n\in\{1,\ldots,\}|\text{if }\{\alpha_i\}_{i\in\{1,\ldots,n\}}\text{ is a family in }F\text{ and }\{x_i\}_{i\in\{1,\ldots,n\}}\text{ is a family in }X\text{ then }L(\sum_{i=1}^n\alpha_i\cdot x_i)=\sum_{i=1}^n\alpha_i\cdot L(x_i)\}$, we have then*
   a. *If $n=1$ then $L(\sum_{i=1}^1\alpha_i\cdot x_i)=L(\alpha_1\cdot x_1)=\alpha_1\cdot L(x_1)=\sum_{i=1}^n\alpha_i\cdot L(x_i)$ and thus $1\in S$*
   b. *If $n\in S$ then if $\{\alpha_i\}_{i\in\{1,\ldots,n+1\}}$ is a field in F and $\{x_i\}_{i\in\{1,\ldots,n+!\}}$ is a field in X then we have $L(\sum_{i=1}^{n+1}\alpha_i\cdot x_i)=L(\sum_{i=1}^n\alpha_i\cdot x_i+\alpha_{n+1}\cdot x_{n+1})=L(\sum_{i=1}^n\alpha_i\cdot x_i)+\alpha_{n+1}\cdot L(x_{n+1})\underset{n\in S}{=}\sum_{i=1}^n\alpha_i\cdot L(x_i)+\alpha_{n+1}\cdot L(x_{n+1})=\sum_{i=1}^{n+1}\alpha_i\cdot L(x_i)$ and thus $n+1\in S$*

   by mathematical induction we have then $S=\{1,\ldots,n\}=\mathbb{N}$ proving (1).
2. *As $\text{support}(\{\alpha_i\cdot x_i\}_{i\in I})\subseteq\text{support}(\{\alpha_i\}_{i\in I})$ (see 10.126). Now if $i\in\text{support}(\{\alpha_i\cdot L(x_i)\}_{i\in I})\Rightarrow\alpha_i\cdot L(x_i)\neq 0\Rightarrow\alpha_i\neq 0\Rightarrow i\in\text{support}(\{\alpha_i\}_{i\in I})$, so $\text{support}(\{\alpha_i\cdot x_i\}_{i\in I}),\text{support}(\{\alpha_i\cdot L(x_i)\}_{i\in I})\subseteq\text{support}(\{\alpha_i\}_{i\in I})$ proving that $\text{support}(\{\alpha_i\cdot x_i\}_{i\in I}),\text{support}(\{\alpha_i\cdot L(x_i)\}_{i\in I})$ is finite and thus that both sums in (2) are well defined. If now $b:\{0,\ldots,n\}\to\text{support}(\{\alpha_i\}_{i\in I})$ then we have $L(\sum_{i\in I}\alpha_i\cdot x_i)=L(\sum_{i=0}^n\alpha_{b_i}\cdot x_{b_i})=L(\sum_{i=1}^{n+1}\alpha_{b_{i-1}}\cdot x_{b_{i-1}})\underset{(1)}{=}\sum_{i=1}^{n+1}\alpha_{b_{i-1}}\cdot L(x_{b_{i-1}})=\sum_{i=0}^n\alpha_{b_i}\cdot L(x_{b_i})=\sum_{i\in I}\alpha_i\cdot L(x_i)$*
3. *This follows from applying (2) to the families $\{\alpha_a\}_{\alpha\in A}$ and $\{a\}_{\alpha\in A}$* □

Example 10.175. If $\{x_i\}_{i\in I}$ is a family with finite support in $\mathbb{C}$ then as $\text{Im}:\mathbb{C}\to\mathbb{R}$ defined by $x\to\text{Im}(x)$ and $\text{Re}:\mathbb{C}\to\mathbb{R}$ defined by $x\to\text{Re}(x)$ are trivially linear we have that $\text{Im}(\sum_{i\in I}x_i)=\sum_{i\in I}\text{Im}(x_i)$ and $\text{Re}(\sum_{i\in I}x_i)=\sum_{i\in I}\text{Re}(x)$

Theorem 10.176. *Let X be a vector space over a field F and $x\in X$ with $x\neq 0$ then there exists a linear mapping $L: X\to F$ such that $L(x)=1$*

Proof. If $x\in X$ with $x\neq 0$ then we have that $\{x\}$ is linearly independent (see 10.107) and thus by 10.157 (and as $x\neq 0$ we have that X is not trivial) there exists a set B with $x\in B$ such that $\{b\}_{b\in B}$ is a basis of X. Define now $L: X\to F$ as follows, if $y\in X$ then there exists a unique $\{\alpha_b^{(y)}\}_{b\in B}$ with finite support such that $y=\sum_{b\in B}\alpha_b^{(y)}\cdot b$ we define then that $L(y)=\sum_{b\in B}\alpha_b^{(y)}$. We prove now that L satisfies the requirement of the theorem.

1. **$(L(x)=1)$** As $x\in B$ we can write $x=1\cdot x$ so that $\{\alpha_b^{(x)}\}_{b\in B}$ is defined by $\alpha_x^{(x)}=1$ and if $b\neq x$ then $\alpha_b^{(x)}=0$ so that $L(x)=\sum_{b\in B}\alpha_b^{(x)}=\alpha_x^{(x)}=1$

2. **(L is linear)** If $y_1, y_2 \in X$ and $\alpha, \beta \in F$ then if $y_1 = \sum_{b \in B} \alpha_b^{(y_1)} \cdot b$, $y_2 = \sum_{b \in B} \alpha_b^{(y_2)} \cdot b$ we have $\alpha \cdot y_1 + \beta \cdot y_2 = \alpha \cdot \left(\sum_{b \in B} \alpha_b^{(y_1)}\right) + \beta \cdot \left(\sum_{b \in B} \alpha_b^{(y_2)}\right) = \sum_{b \in B} \left(\alpha \cdot \alpha_b^{(y_1)} + \beta \cdot \alpha_b^{(y_2)}\right) \cdot b$ so that $L(\alpha \cdot y_1 + \beta \cdot y_2) = L\left(\sum_{b \in B} \left(\alpha \cdot \alpha_b^{(y_1)} + \beta \cdot \alpha_b^{(y_2)}\right) \cdot b\right) = \sum_{b \in B} \left(\alpha \cdot \alpha_b^{(y_1)} + \beta \cdot \alpha_b^{(y_2)}\right) = \alpha \cdot \sum_{b \in B} \alpha_b^{(y_1)} + \beta \cdot \sum_{b \in B} \alpha_b^{(y_2)} = \alpha \cdot L(y_1) + \beta \cdot L(y_2)$ □

We can use the above theorem to prove the following handy theorem

Theorem 10.177. *If X is a vector space over a field F with characterization zero, $x, y \in X$ and if for every linear mapping $L: X \rightarrow F$ we have $L(x) = L(y)$ then we must have $x = y$*

Proof. Assume that $x \neq y \Rightarrow x - y \neq 0$ then using the previous theorem we can find a linear mapping L such that $L(x - y) = 1$ but then we have $1 = L(x - y) = L(x) - L(y) = 0$ giving the contradiction $1 = 0$. □

Theorem 10.178. *If X is a vector space over a field F and $\{e_i\}_{i \in \{1,\ldots,n\}}$ a linear independent set then there exists a linear mapping $L: X \rightarrow F$ such that $L(e_1) = 1$ and $L(e_i) = 0, i \in \{2, \ldots, k\}$.*

Proof. As $\{e_i\}_{i \in \{1,\ldots,n\}}$ is linear independent it can be extended by 10.157 to a basis $\{f_i\}_{i \in \{1,\ldots,k\}}$ of X (so that $\forall i \in \{1, \ldots, n\}$ we have $e_i = f_i$. Given now $y \in X$ where $y = \sum_{i \in \{1,\ldots,k\}} y_i \cdot f_i$ is the unique expansion of y in the basis $\{f_i\}_{i \in \{1,\ldots,k\}}$. Define then $L(y) = y_1$, this gives us the function (because the expansion is unique) that we are searching for.

1. **($y = e_1$)** then $y = e_1 = f_1 = 1 \cdot f_1$ or $y_1 = 1$ so that $L(y) = 1$
2. **($y = e_i$ where $i \in \{2, \ldots, n\}$)** then $y = e_i = 1 \cdot e_i$ or $y_1 = 0$ so that $L(e_i) = 0$
3. **(L is linear)** If $x = \sum_{i \in \{1,\ldots,k\}} x_i$, $y = \sum_{i \in \{1,\ldots,k\}} y_i \in X$ and $\alpha, \beta \in F$ then $\alpha \cdot x + \beta \cdot y = \alpha \cdot \sum_{i \in \{1,\ldots,k\}} x_i + \beta \cdot \sum_{i \in \{1,\ldots,k\}} y_i = \sum (\alpha \cdot x_i + \beta \cdot y_i)$ so that $L(\alpha \cdot x + \beta \cdot y) = \alpha \cdot x_1 + \beta \cdot y_1 = \alpha \cdot L(x) + \beta \cdot L(y)$ □

Definition 10.179. *If X, Y are vector spaces over a field F then a function L: $X \rightarrow Y$ is a* ***linear isomorphism*** *iff*

1. *L is a bijection*
2. *L is linear*

Example 10.180. $\mathcal{C}: \langle \mathbb{R}^2, +, \cdot \rangle \rightarrow \langle \mathbb{C}, +, \cdot \rangle$ defined by $\mathcal{C}((x, y)) = x + i \cdot y$ is a isomorphism (here we consider $\langle \mathbb{C}, +, \cdot \rangle$ a real vector space (see 10.112)

Proof. Using 9.39 we have that $\mathcal{C}$ is a bijection so we have only to prove linearity. So let $\alpha, \beta \in \mathbb{R}$ and $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$ then $\mathcal{C}(\alpha \cdot (x_1, y_1) + \beta \cdot (x_2, y_2)) = \mathcal{C}((\alpha \cdot x_1 + \beta \cdot x_2, \alpha \cdot y_1 + \beta \cdot y_2)) = \alpha \cdot x_1 + \beta \cdot y_1 + i \cdot (\alpha \cdot y_1 + \beta \cdot y_2) = \alpha \cdot (x_1 + i \cdot y_1) + \beta \cdot (x_2 + i \cdot y_2) = \alpha \cdot \mathcal{C}((x_1, y_1)) + \beta \cdot \mathcal{C}((x_2, y_2))$ □

Theorem 10.181. *If X, Y are vector spaces over a field F and $L: X \rightarrow Y$ a linear isomorphism then $L^{-1}: Y \rightarrow X$ is defined and is also a linear isomorphism.*

Proof. Using 2.38 we have that $L^{-1}: Y \rightarrow X$ is defined and is also a bijection. If α, $\beta \in F$ and $x, y \in X$ then there exists a $x' = L^{-1}(x), y' = L^{-1}(y)$ such that $L(x') = x$, $L(y') = y$. Now $L^{-1}(\alpha \cdot x + \beta \cdot y) = L^{-1}(\alpha \cdot L(x') + \beta \cdot L(y')) = L^{-1}(L(\alpha \cdot x' + \beta \cdot y')) = \alpha \cdot x' + \beta \cdot y' = \alpha \cdot L^{-1}(x) + \beta \cdot L^{-1}(y)$ proving the linearity of L^{-1}. $\square$

Theorem 10.182. *If X, Y, Z are vector spaces over a field F and $L_1: X \rightarrow Y$, $L_2: Y \rightarrow Z$ linear mappings (linear isomorphism) then $L_2 \circ L_1: X \rightarrow Z$ is a linear mapping (linear isomorphism).*

Proof. First if $\alpha, \beta \in F$ and $x, y \in X$ then $(L_2 \circ L_1)(\alpha \cdot x + \beta \cdot y) = L_2(L_1(\alpha \cdot x + \beta \cdot y)) = L_2(\alpha \cdot L(x) + \beta \cdot L(y)) = \alpha \cdot L_2(L_1(x)) + \beta \cdot L_2(L_1(y)) = \alpha \cdot (L_2 \circ L_1)(x) + \beta \cdot (L_2 \circ L_1)(y)$.

Second if L_1, L_2 are linear isomorphism they are linear and thus $L_2 \circ L_1$ is linear (see above) (and a bijection (see 2.46). $\square$

Theorem 10.183. *Let $\langle X, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$, Y a set and $L: X \rightarrow Y$ a bijection. define then $+_L: Y \times Y \rightarrow Y$ by $(x, y) \rightarrow x +_L y = L(L^{-1}(x) + L^{-1}(y))$ and $\cdot_L: F \times Y \rightarrow Y$ by $(\alpha, x) \rightarrow \alpha \cdot_L x = L(\alpha \cdot L^{-1}(x))$. Then $\langle Y, +_L, \cdot_L \rangle$ forms a vector space over $\langle F, +, \cdot \rangle$ with zero element $L(0)$ and $L: X \rightarrow Y$ is then a isomorphism. Further if $\{e_i\}_{i \in I}$ is a basis in X then $\{L(e_i)\}_{i \in I}$ is a basis in $\langle Y, +_L, \cdot_L \rangle$*

Proof. First we prove that $\langle Y, +_L \rangle$ is a abelian group

1. **(associativity)** $x +_L (y +_L z) = x +_L (L(L^{-1}(y) + L^{-1}(z))) = L(L^{-1}(x) + L^{-1}(L(L^{-1}(y) + L^{-1}(z)))) = L(L^{-1}(x) + (L^{-1}(y) + L^{-1}(z))) = L((L^{-1}(x) + L^{-1}(y)) + L^{-1}(z)) = L(L^{-1}(L(L^{-1}(x) + L^{-1}(y))) + L^{-1}(z)) = L(L^{-1}(x +_L y) + L^{-1}(z)) = ((x +_L y) +_L z)$
2. **(commutativity)** $x +_L y = L(L^{-1}(x) + L^{-1}(y)) = L(L^{-1}(y) + L^{-1}(x)) = y +_L x$
3. **(neutral element)** $x +_L L(0) = L(L^{-1}(x) + L^{-1}(L(0))) = L(L^{-1}(x) + 0) = L(L^{-1}(x)) = x$ so $x +_L L(0) = x \underset{\text{commutativity}}{=} L(0) +_L x$

Next we prove the remaining requirements of a vector space

1. $\alpha \cdot_L (x +_L y) = \alpha \cdot_L (L(L^{-1}(x) + L^{-1}(y))) = L(\alpha \cdot L^{-1}(L(L^{-1}(x) + L^{-1}(y)))) = L(\alpha \cdot (L^{-1}(x) + L^{-1}(y))) = L(\alpha \cdot L^{-1}(x) + \alpha \cdot L^{-1}(y)) = L(L^{-1}(L(\alpha \cdot L^{-1}(x))) + L^{-1}(L(\alpha \cdot L^{-1}(y)))) = L(\alpha \cdot L^{-1}(x)) +_L L(\alpha \cdot L^{-1}(y)) = \alpha \cdot_L x +_L \alpha \cdot_L y$
2. $(\alpha + \beta) \cdot_L x = L((\alpha + \beta) \cdot L^{-1}(x)) = L(\alpha \cdot L^{-1}(x) + \beta \cdot L^{-1}(x)) = L(L^{-1}(L(\alpha \cdot L^{-1}(x))) + L^{-1}(L(\beta \cdot L^{-1}(x)))) = L(\alpha \cdot L^{-1}(x)) +_L L(\beta \cdot L^{-1}(x)) = \alpha \cdot_L x + \beta \cdot_L x$
3. $(\alpha \cdot \beta) \cdot_L x = L((\alpha \cdot \beta) \cdot L^{-1}(x)) = L(\alpha \cdot (\beta \cdot L^{-1}(x))) = L(\alpha \cdot L^{-1}(L(\beta \cdot L^{-1}(x)))) = \alpha \cdot_L L(\beta \cdot L^{-1}(x)) = \alpha \cdot_L (\beta \cdot_L x)$

Next we prove that L is linear (and as it is bijective) thus a isomorphism

$L(\alpha \cdot x + \beta \cdot y) = L(L^{-1}(L(\alpha \cdot x)) + L^{-1}(L(\beta \cdot y))) = L(\alpha \cdot x) +_L L(\beta \cdot y) = L(\alpha \cdot L^{-1}(L(x))) + L(\beta \cdot L^{-1}(L(y))) = \alpha \cdot_L L(x) + \beta \cdot_L L(y)$

Finally if $\{e_i\}_{i\in I}$ is a basis for $\langle X,+,\cdot\rangle$ then we have for $\{L(e_i)\}_{i\in I}$

1. If $\{\alpha_i\}_{i\in I}$ is a family with finite support such that $0=L(0)=\sum_{i\in I}\alpha_i\cdot L(e_i) \underset{L\,\text{is linear}\,10.174}{=} L(\sum_{i\in I}\alpha_i\cdot e_i) \underset{L\,\text{is bijective}}{\Rightarrow} 0=\sum_{i\in I}\alpha_i\cdot e_i \underset{\{e_i\}_{i\in I}\,\text{is linear independent}}{\Rightarrow} \forall i\in I$ we have $\alpha_i=0$ proving that
$$\{L(e_i)\}_{i\in I}\,\text{is independent} \tag{10.32}$$
2. If $y\in Y$ then there exists a $x\in X$ such that $y=L(x)$, as $\{e_i\}_{i\in I}$ is a basis there exists a $\{\alpha_v\}_{v\in\{e_i|i\in I\}}$ with finite support such that $x=\sum_{v\in\{e_i|i\in I\}}\alpha_v\cdot v$ and thus $y=L(x)=L\big(\sum_{v\in\{e_i|i\in I\}}\alpha_v\cdot v\big)=\sum_{v\in\{e_i|i\in I\}}\alpha_v\cdot L(v)$ so
$$y=\sum_{v\in\{e_i|i\in I\}}\alpha_v\cdot L(v) \tag{10.33}$$
Now as $L: X\rightarrow Y$ is a bijection we have that $L_{|\{e_i|i\in I\}}: \{e_i|i\in I\}\rightarrow L(\{e_i|i\in I\})=\{L(e_i)|i\in I\}$ is a bijection, now define $\{\beta_v\}_{v\in\{L(e_i)|i\in I\}}$ by $\beta_v=\alpha_{L^{-1}(v)}$ then we have $\beta_{L(v)}=\alpha_{L^{-1}(L(v))}=\alpha_v$ and thus using 10.33 we have $y=\sum_{v\in\{e_i|i\in I\}}\beta_{L(v)}\cdot L(v)=\sum_{v\in\{e_i|i\in I\}}\beta_{L_{|\{e_i|i\in I\}}(v)}\cdot L_{|\{e_i|i\in I\}}(v) \underset{10.44}{=} \sum_{v\in\{L(e_i)|i\in I|\}}\beta_v\cdot v$ and thus $y\in\mathcal{S}(\{L(e_i)|i\in I|\})$ and this proves that:
$$Y=\mathcal{S}(\{L(e_i)|i\in I\}) \tag{10.34}$$

Using 10.32 and 10.34 we have then that $\{L(e_i)\}_{i\in I}$ is linear independent. □

Definition 10.184. *Let X,Y be vector spaces over a field F then* $\mathrm{Hom}(X,Y)=\{L\in M(X,Y)|L\,\text{is a linear mapping}\}$ *(for a definition of $M(X,Y)$ see 10.113). [In other words* $\mathrm{Hom}(X,Y)$ *is the set of all linear mappings from X to Y]*

Theorem 10.185. *Let X,Y be a vector space over a field F then* $\mathrm{Hom}(X,Y)$ *is a non empty subspace of $M(X,Y)$ (see 10.113 for definition of $M(X,Y)$) and thus by 10.109* $\mathrm{Hom}(X,Y)$ *is a vector space.*

Proof. First if we define $0: X\rightarrow Y$ by $x\rightarrow 0(x)=0$ then as $0(\alpha\cdot x+\beta\cdot y)=0=0+0=\alpha\cdot 0(x)+\beta\cdot 0(y)$ so $0\in\mathrm{Hom}(X,Y)\Rightarrow\emptyset\neq\mathrm{Hom}(X,Y)$.

Finally if $\alpha,\beta\in F$, $L_1,L_2\in\mathrm{Hom}(X,Y)$ then if $\alpha',\beta'\in F$ and $x,y\in X$ we have $(\alpha\cdot L_1+\beta\cdot L_2)(\alpha'\cdot x+\beta'\cdot y)=\alpha\cdot L_1(\alpha'\cdot x+\beta'\cdot y)+\beta\cdot L_2(\alpha'\cdot x+\beta'\cdot y)=\alpha\cdot(\alpha'\cdot L_1(x))+\alpha\cdot(\beta'\cdot L_1(y))+\beta\cdot(\alpha'\cdot L_2(x))+\beta\cdot(\beta'\cdot L_2(y))=\alpha'\cdot(\alpha\cdot L_1(x)+\beta\cdot L_2(x))+\beta'\cdot(\alpha\cdot L_1(y)+\beta\cdot L_2(y))=\alpha'\cdot(\alpha\cdot L_1+\beta\cdot L_2)(x)+\beta'(\alpha\cdot L_1+\beta\cdot L_2)(y)$ proving that $\alpha\cdot L_1+\beta\cdot L_2$ is linear and thus that $\alpha\cdot L_1+\beta\cdot L_2\in\mathrm{Hom}(X,Y)$. □

Definition 10.186. *Let V be a vector space over a field F (which is itself a vector space by 10.110) then we note the vector space* $\mathrm{Hom}(V,F)$ *as V^* and call it the* ***dual space*** *of V*

Theorem 10.187. *If V is a finite vector space over the field F with a basis $\{e_i\}_{i\in\{1,\ldots,m\}}$ (so* $\dim(V)=m$*). Then V^* is finite dimensional with basis $\{e_i^*\}_{i\in\{1,\ldots,m\}}$ where $\forall i\in\{1,\ldots,m\}$ we define $e_i^*: V\rightarrow F$ by $e_i^*\big(\sum_{i\in\{1,\ldots,m\}}\lambda_i\cdot e_i\big)=\lambda_i$*

Proof. First we have to prove that e_i^* is linear so if we take $x = \sum_{i \in \{1,\ldots,m\}} \lambda_i \cdot e_i$, $y = \sum_{i \in \{1,\ldots,m\}} \lambda_i' \cdot e_i$ and $\alpha, \beta \in F$ then $\alpha \cdot x + \beta \cdot y = \alpha \cdot \left(\sum_{i \in \{1,\ldots,m\}} \lambda_i \cdot e_i\right) + \beta\left(\sum_{i \in \{1,\ldots,m\}} \lambda_i' \cdot e_i\right) = \sum_{i \in \{1,\ldots,m\}} (\alpha \cdot \lambda_i + \beta \cdot \lambda_i')$ so that $e_i^*(\alpha \cdot x + \beta \cdot y) = \alpha \cdot \lambda_i + \beta \cdot \lambda_i' = \alpha \cdot e_i^*(x) + \beta \cdot e_i^*(y)$ proving linearity.

Second we prove that $\mathcal{S}(\{e_i^*\}_{i \in \{1,\ldots,m\}}) = V^*$, if $f \in V^*$ then $f: V \to F$ is a linear mapping. Define now $\{\alpha_i\}_{i \in \{1,\ldots,m\}}$ by $\alpha_i = f(e_i)$ then $\forall x \in V$ we have $x = \sum_{i \in \{1,\ldots,m\}} \lambda_i \cdot e_i \Rightarrow f(x) = \sum_{i \in \{1,\ldots,m\}} \lambda_i \cdot f(e_i) = \sum_{i \in \{1,\ldots,m\}} \alpha_i \cdot \lambda_i = \sum_{i \in \{1,\ldots,m\}} \alpha_i \cdot e_i^*(x) = \left(\sum_{i \in \{1,\ldots,m\}} \alpha_i \cdot e_i^*\right)(x)$ proving that $f = \sum_{i \in \{1,\ldots,m\}} \alpha_i \cdot e_i^*$ and thus that $V^* \subseteq \mathcal{S}(\{e^*_{i \in \{1,\ldots,m\}}\})$ which together with the trivial $\mathcal{S}(\{e_i^*\}_{i \in \{1,\ldots,m\}}) \subseteq V^*$ we have $V^* = \mathcal{S}(\{e_i^*\}_{i \in \{1,\ldots,m\}})$.

Finally to prove that $\{e_i\}_{i \in \{1,\ldots,m\}}$ is linear independent assume that $\sum_{i \in \{1,\ldots,m\}} \alpha_i \cdot e_i^* = 0$ (the mapping $V \to F$ with $0(x) = 0$) then $\forall i \in \{1, \ldots, m\}$ we have that $0 = 0(e_i) = \left(\sum_{j \in \{1,\ldots,m\}} \alpha_j \cdot e_j^*(e_i)\right) = \sum_{j \in \{1,\ldots,m\}} \alpha_j \cdot \delta_{i,j} = \alpha_i \Rightarrow \alpha_i = 0$ proving linear independence. □

Definition 10.188. *If X is a vector space over a field F then* $\mathrm{GL}(X) = \{L \in \mathrm{Hom}(X, X) | L \text{ is a bijection}\} \subseteq \mathrm{Hom}(X, X)$. *So* $\mathrm{GL}(X)$ *is the set of all the linear isomorphism's from $X \to X$.*

Theorem 10.189. *If X is a vector space over a field F then $\langle \mathrm{GL}(X), \circ \rangle$ is a group with neutral element 1_X (here $\circ$ is the function composition operator). This group is called the **general linear group**. Furthermore if we defined $\triangleright: \mathrm{GL}(X) \times X \to X$ by $(L, x) \to \triangleright(L, x) \underset{\text{notation}}{=} L \triangleright x \underset{\text{defined}}{=} L(x)$ then $\triangleright$ is a **faithful left action** (see 3.19, 3.16)*

Proof. First we must prove that $\circ: \mathrm{GL}(X) \times \mathrm{GL}(X) \to \mathrm{GL}(X)$ is indeed a function, essential we must prove that if $L_1, L_2 \in \mathrm{GL}(X)$ then that $L_1 \circ L_2 \in \mathrm{GL}(X)$ this follows directly from 10.182. Now for the rest of the group axioms:

1. **(associativity)** This follows from the associativity of the composition of functions (see 1.54)
2. **(neutral element)** If $L \in \mathrm{GL}(X)$ then trivially we have $1_X \in \mathrm{GL}(X)$ and $L \circ 1_X = L = 1_X \circ L$
3. **(inverse element)** If $L \in \mathrm{GL}(X)$ then using 10.181 we have that $L^{-1} \in \mathrm{GL}(X)$ and trivially we have $L \circ L^{-1} = 1_X = L^{-1} \circ L$

To prove that $\triangleright$ is a left action note that:

1. $\forall x \in X$ we have $1_X \triangleright x = 1_X(x) = x \Rightarrow 1_X \triangleright x = x$
2. $\forall L, L' \in \mathrm{GL}(X)$ and $\forall x \in X$ we have $(L \circ L') \triangleright x = (L \circ L')(x) = L(L'(x)) = L \triangleright (L'(x)) = L \triangleright (L' \triangleright x)$

To prove that the left action is faithful, if $L \in \mathrm{GL}(X)$ take then $L_\triangleright: X \to X$ defined by $x \to L_\triangleright(x) = L \triangleright x = L(x)$, we have then the following:

1. If $L_\triangleright = 1_X$ then $\forall x \in X$ we have $L(x) = L_\triangleright(x) = 1_X(x) = x \Rightarrow L = 1_X$ the neutral element in $\mathrm{GL}(X)$
2. If $L = 1_X$ then $\forall x \in X$ we have $L_\triangleright(x) = L(x) = 1_X(x) = x \Rightarrow L_\triangleright = 1_X$ □

Theorem 10.190. *Let X be a vector space over a field F and let $\emptyset \neq Y \subseteq X$ be a subspace of X then $\sim_Y$ defined by $\sim_Y = \{(x, y) \in X \times X | x - y \in Y\}$ is a equivalence relation.*

Proof. First as $\emptyset \neq Y$ we have that Y is a vector space over the field F and thus if 0 is the neutral element in X then $0 \in Y$. Next we prove that $\sim_X$ is a equivalence relation

1. **(reflexive)** If $x \in X$ then $x - x = 0 \in Y \Rightarrow x \sim_Y x$
2. **(symmetric)** If $x, y \in X$ and $x \sim_Y y$ then $x - y \in Y \Rightarrow (-1) \cdot (x - y) \in Y \Rightarrow y - x \in Y \Rightarrow y \sim_Y x$
3. **(transitive)** If $x, y, z \in X$ and $x \sim_Y, y \sim_Y z$ then we have $x - y \in Y$ and $y - z \in Y$ and thus $x - z = (x - y) + (y - z) \in Y \Rightarrow x \sim_Y z$ □

Theorem 10.191. *Let X be a vector space over a field F and $\emptyset \neq Y \subseteq X$ a subspace of X then if we define $X/Y = X/\sim_Y = \{\sim_Y[x] | x \in X\}$ and*

1. *$+: X/Y \times X/Y \to X/Y$ defined by $(\sim_Y[x], \sim_Y[y]) \to \sim_Y[x] + \sim_Y[y] = \sim_Y[x + y]$*
2. *$\cdot: F \times X/Y \to X/Y$ defined by $(\alpha, \sim_Y[x]) \to \alpha \cdot \sim_Y[x] = \sim_Y[\alpha \cdot x]$*

*then these are well defined functions and $\langle X/Y, +, \cdot \rangle$ forms a vector space over F with neutral element $Y = \sim_Y[0]$, this vector space is called the **factor space**. Also the canonical surjective function (see 2.127) $\pi_Y: X \to X/Y$ defined by $x \to \pi_Y(x) = \sim_Y[x]$ is a linear mapping.*

Proof. First we prove that the operations are well defined. If $\sim_Y[x] = \sim_Y[x']$ and $\sim_Y[y] = \sim_Y[y'] \Rightarrow x - x' \in Y$ and $y - y' \in Y$ and we have:

1. $(x + y) - (x' + y') = (x - x') + (y - y') \in Y \Rightarrow (x + y) \sim_Y (x' + y')$ so we have $\sim_Y[x + y] = \sim_Y[x' + y']$ proving that $+: X/Y \times X/Y \to X/Y$ is well defined.
2. $\alpha \cdot x - \alpha \cdot x' = \alpha \cdot (x - x') \in Y \Rightarrow \alpha \cdot x \sim_Y \alpha \cdot x'$ so we have $\sim[\alpha \cdot x] = \sim[\alpha \cdot x']$ proving that $\cdot: F \times X/Y \to X/Y$ is well defined.

Next we prove that $\langle X/Y, +, \cdot \rangle$ is a vector space over the field F.

1. **(Abelian group axioms)** Let $\sim_Y[x], \sim_Y[y], \sim_Y[z] \in X/Y$
   a. **(associativity)** $\sim_Y[x] + (\sim_Y[y] + \sim_Y[z]) = \sim_Y[x] + \sim_Y[y + z] = \sim_Y[x + (y + z)] = \sim_Y[(x + y) + z] = \sim_Y[x + y] + \sim_Y[z] = (\sim_Y[x] + \sim_Y[y]) + \sim_Y[z]$
   b. **(commutativity)** $\sim_Y[x] + \sim_Y[y] = \sim_Y[x + y] = \sim_Y[y + x] = \sim_Y[y] + \sim_Y[x]$
   c. **(neutral element)** As $Y \neq \emptyset$ we have that Y is a a subspace and thus $0 \in Y$, if $y \in Y \Rightarrow y = y - 0 \in Y \Rightarrow y \sim_Y 0 \Rightarrow y \in \sim_Y[0] \Rightarrow Y \subseteq \sim_Y[0]$, if $y \in \sim_Y[0]$ then $y = y - 0 \in Y \Rightarrow \sim_Y[0] \subseteq Y$ so we have
   $$\sim_Y[0] = Y \tag{10.35}$$
   Then $\sim_Y[x] + Y = \sim_Y[x] + \sim_Y[0] = \sim_Y[x + 0] = \sim_Y[x] \Rightarrow x = \sim_Y[x] + Y \underset{\text{commutativity}}{=} Y + \sim_Y[x]$
   d. **(inverse element)** $\sim_Y[x] + \sim_Y[-x] = \sim_Y[x - x] = \sim_Y[0] = Y \Rightarrow Y = \sim_Y[x] + \sim_Y[-x] \underset{\text{commutativity}}{=} \sim_Y[-x] + \sim_Y[x]$

2. **(Vector space axioms)**

   a. Let $\alpha \in F$ and $\sim_Y[x], \sim_Y[y] \in X/Y$ then $\alpha \cdot (\sim_Y[x] + \sim_Y[y]) = \alpha \cdot \sim_Y[x+y] = \sim_Y[\alpha \cdot (x+y)] = \sim_Y[\alpha \cdot x + \alpha \cdot y] = \sim_Y[\alpha \cdot x] + \sim_Y[\alpha \cdot y] = \alpha \cdot \sim_Y[x] + \alpha \cdot \sim_Y[y]$

   b. Let $\alpha, \beta \in F$ and $\sim_Y[x] \in X/Y$ then $(\alpha + \beta) \cdot \sim_Y[x] = \sim_Y[(\alpha + \beta) \cdot x] = \sim_Y[\alpha \cdot x + \beta \cdot x] = \sim_Y[\alpha \cdot x] + \sim_Y[\beta \cdot x] = \alpha \cdot \sim_Y[x] + \beta \cdot \sim_Y[x]$

   c. Let $\alpha, \beta \in F$ and $\sim_Y[x] \in X/Y$ then $(\alpha \cdot \beta) \cdot \sim_Y[x] = \sim_Y[(\alpha \cdot \beta) \cdot x] = \sim_Y[\alpha \cdot (\beta \cdot x)] = \alpha \cdot \sim_Y[\beta \cdot x] = \alpha \cdot (\beta \cdot \sim_Y[x])$

   d. If $\sim_Y[x] \in X/Y$ then we have $1 \cdot \sim_Y[x] = \sim_Y[1 \cdot x] = \sim_Y[x]$

Finally if $\alpha, \beta \in F$ and $x, y \in X$ then $\pi_Y(\alpha \cdot x + \beta \cdot y) = \sim_Y[\alpha \cdot x + \beta \cdot y] = \sim_Y[\alpha \cdot x] + \sim_Y[\beta \cdot y] = \alpha \cdot \sim_Y[x] + \beta \cdot \sim_Y[y] = \alpha \cdot \pi_Y(x) + \beta \cdot \pi_Y(y)$ proving the linearity of π_Y. □

Definition 10.192. (Direct Internal Sum) *If V is a vector space over a field F then $V = X \oplus Y$ iff X, Y are non empty sub spaces of V, $X \bigcap Y = \{0\}$ and $V = \{x + y | (x, y) \in X \times Y\}$*

Theorem 10.193. *If V is a vector space over a field F then if $V = X \oplus Y$ then $\forall a \in V$ we have that there exists* ***unique*** *$x \in X, y \in Y$ such that $a = x + y$*

Proof. If $a \in V = X \oplus Y$ then there exists a $x \in X, y \in Y$ such that $a = x + y$. Assume now that also $a = x' + y'$ (where $x' \in X$, $y' \in Y$) then $0 = a - a = (x - x') + (y - y') \Rightarrow x - x' = y' - y \Rightarrow X \ni x - x' = y' - y \in Y \Rightarrow (x - x'), (y' - y) \in X \bigcap Y = \{0\} \Rightarrow x = x'$, $y = y'$ □

Theorem 10.194. *If V is a vector space over a field F and $\emptyset \neq X \subseteq V$ is a sub space of V then there exist a sub space $\emptyset \neq Y \subseteq V$ such that $V = X \oplus Y$*

Proof. As $X \neq \emptyset$ X is a vector space and thus $0 \in X$ we have then the following cases to consider for X

1. $(\boldsymbol{X = V})$ take then $Y = \{0\}$ then $X \bigcap Y = V \bigcap \{0\} = \{0\}$ and if $v \in V \Rightarrow v = v + 0$ where $(v, 0) \in V \times \{0\} = X \times Y \Rightarrow V = X \oplus Y$
2. $(\boldsymbol{X = \{0\}})$ take then $Y = V$ then $X \bigcap Y = \{0\} \bigcap Y = \{0\}$ and if $v \in V \Rightarrow v = 0 + v$ where $(0, v) \in \{0\} \times V = X \times Y \Rightarrow V = X \oplus Y$
3. $(\boldsymbol{X \neq \{0\}, V})$ Then as $X \neq \{0\}$ we have that X is not a trivial vector space. Using 10.155

we have the existence of a $B_X \subseteq X$ such that $\{b\}_{b \in B_X}$ is a basis for X. Using the linear independence of B_X and 10.157 we find a B with $B_X \subseteq B$ and $\{b\}_{b \in B}$ is a basis for V. If now $B_X = B$ then $X = \mathcal{S}(B_X) = \mathcal{S}(B) = V$ contradicting $X \neq V$ so we must have $B_X \subset B$ and thus have $\emptyset \neq B | B_X$. Using 10.130 we have then that $\mathcal{S}(B \setminus B_X)$ is a vector space over F (and also a subspace of V), take then $Y = \mathcal{S}(B \setminus B_X)$. If now $v \in V$ then as $\{b\}_{b \in B}$ is a basis of V there exists a $\{\alpha_b\}_{b \in B}$ with finite support such that $v = \sum_{b \in B} \alpha_b \cdot b = \sum_{b \in B_X} \alpha_b \cdot b + \sum_{b \in B \setminus B_x} \alpha_b \cdot b = x + y$ where $x = \sum_{b \in B_X} \alpha_b \cdot b \in \mathcal{S}(B_X) = X$ and $y = \sum_{b \in B \setminus B_X} \alpha_b \cdot b \in \mathcal{S}(B \setminus B_X) = Y$ also if $x \in X \subseteq V, y \in Y \subseteq V$ we have $x + y \in V$. So we have

$$v \in V \Leftrightarrow \exists (x, y) \in X \times Y \text{ such that } v = x + y \tag{10.36}$$

If now $z \in X \bigcap Y$ then there exists families $\{\beta_b\}_{b \in B_X}$, $\{\tau_b\}_{b \in B \setminus B_X}$ in F with finite support such that $\sum_{b \in B_X} \beta_b \cdot b = z = \sum_{b \in B \setminus B_X} \tau_b \cdot b \Rightarrow \sum_{b \in B_X} \beta_b \cdot b - \sum_{b \in B \setminus B_X} \tau_b \cdot b = 0$. Define now $\{\gamma_b\}_{b \in B}$ by $\gamma_b = \left\{ \begin{smallmatrix} \beta_b \text{ if } b \in B_X \\ -\tau_b \text{ if } b \in B \setminus B_X \end{smallmatrix} \right.$ which, because $\text{support}(\{\gamma_b\}_{b \in B}) \subseteq \text{support}(\{\beta_b\}_{b \in B_X}) \bigcup \text{support}(\{\tau_b\}_{b \in B \setminus B_X})$, has finite support. The $\sum_{b \in B} \gamma_b \cdot b = \sum_{b \in B_X} \gamma_b \cdot b + \sum_{b \in B \setminus B_X} \gamma_b \cdot b = \sum_{b \in B_X} \beta_b \cdot b + \sum_{b \in B \setminus B_X} (-1) \cdot \tau_b \cdot b = \sum_{b \in B_X} \beta_b \cdot b - \sum_{b \in B \setminus B_X} \tau_b \cdot b = 0$. Using the independence of B we have then $\forall b \in B$ we have $\gamma_b = 0$, so $\forall b \in B_X \subseteq B$ we have $\alpha_b = \beta_b = 0$ and thus $z = \sum_{b \in B_X} \beta_b \cdot b = \sum_{b \in B_X} 0 \cdot b = 0$ proving that

$$X \bigcap Y = \{0\} \tag{10.37}$$

From 10.36 and 10.37 we have then $V = X \oplus Y$ □

Theorem 10.195. *If V is a non trivial vector space over a field F with a basis $\{b_i\}_{i \in \{1,\ldots,n\}}$ then if $\{a_i\}_{i \in \{1,\ldots,n+1\}}$ is a finite family in V we have that $\{a_i\}_{i \in \{1,\ldots,n+1\}}$ is linear dependent.*

Proof. We prove this by induction so let $S = \{n \in \{1, \ldots\} | \text{If } V \text{ is } a \text{ vector space with } a \text{ finite basis } \{b_i\}_{i \in \{1,\ldots,m\}}\ m \leqslant n \text{ such that if } \{a_i\}_{i \in \{1,\ldots,n+1\}} \text{ is } a \text{ finite family in } V \text{ then } \{a_i\}_{i \in \{1,\ldots,n\}} \text{ is linear depdendent}\}$ we have then:

1. **($n = 1$)** If V is a vector space with a basis $\{v_i\}_{i \in \{1,\ldots,m\}}$ where $m \leqslant 1$ then as V is not trivial we have by 10.150 that $\{1, \ldots, m\} \neq \emptyset \Rightarrow m = 1$ and thus $\{v_i\}_{i \in \{1\}}$ is a basis of V, If now $A = \{a_i\}_{i \in \{1,\ldots,2\}} \Rightarrow A = \{a_1, a_2\}$ then there exists a $\{\alpha_i\}_{i \in \{1\}}, \{\beta_i\}_{i \in \{1\}}$ such that $a_1 = \sum_{i \in \{1\}} \alpha_i \cdot v_i = \alpha_1 \cdot v_1$ and $\beta_1 = \sum_{i \in \{1\}} \beta_i \cdot v_i = \beta_1 \cdot v_1$. We have then the following possibilities:

   a. **($\alpha_1 = 0 \vee \beta_1 = 0$)** $\Rightarrow a_1 = 0 \vee a_2 = 0 \Rightarrow A$ is linear dependent

   b. **($\alpha_1, \alpha_2 \neq 0$)** define then $\{\gamma_i\}_{i \in \{1,2\}}$ such that $\gamma_{a_1} = 1$ and $\gamma_{a_2} = -\alpha_1 \cdot \beta_1^{-1}$ then $\sum_{i \in \{1,2\}} \gamma_i \cdot a_i = 1 \cdot a_1 - \alpha_1 \cdot \beta_1^{-1} = \alpha_1 \cdot v_1 - \alpha_1 \cdot \beta_1^{-1} (\beta_1 \cdot v_1) = \alpha_1 \cdot v_1 - \alpha_1 \cdot v_1 = 0$ which as $\gamma_{a_1} = 1 \neq 0$ proves that A is linear dependent.

   So in all cases we have that A is linear dependent and thus we have that $1 \in S$.

2. **($n \in S$)** If V is a vector space with a basis $\{b_i\}_{i \in \{1,\ldots,m\}}$ where $m \leqslant n + 1$. As V is non trivial we have by 10.150 that $\{1, \ldots, m\} \neq \emptyset \Rightarrow 1 \leqslant m \leqslant n + 1$. Let now $A = \{a_i\}_{i \in \{1,\ldots,n+2\}}$ a family in V. We have now the following cases:

   a. **($m \leqslant n$)** then as $n \in S$, $m \leqslant n$ and we have that $\{a_i\}_{i \in \{1,\ldots,n+1\}}$ is linear dependent. Finally using 10.147 and $\{a_i\}_{i \in \{1,\ldots,n+1\}} \preccurlyeq \{a_i\}_{i \in \{1,\ldots,n+2\}}$ we have that $A = \{a_i\}_{i \in \{1,\ldots,n+2\}}$ is linear dependent.

   b. **($n < m$)** then we have $n < m \leqslant n + 1 \Rightarrow n + 1 \leqslant m \leqslant n + 1 \Rightarrow m = n + 1$. We have now the following possibilities for a_1, a_{n+2}

      i. **($a_1 = a_{n+2}$)** then we have that A is linear dependent

ii. $(\boldsymbol{a_1 \neq a_{n+2} \wedge (a_{n+2} = 0 \vee a_1 = 0)})$ then A is linear dependent.

iii. $(\boldsymbol{a_1 \neq a_{n+2} \wedge a(n+2), a(1) \neq 0})$. Take then

$$V_1 = V / \mathcal{S}(\{a_{n+2}\}) \tag{10.38}$$

and the surjection $\pi_{\mathcal{S}(\{a_{n+2}\})}: V \rightarrow V_1$. Define now $\{\bar{b_i}\}_{i \in \{1, \ldots, n+1\}}$ where $\forall i \in \{1, \ldots, n+1\}$ we have $\bar{b_i} = \pi_{\mathcal{S}(\{a_{n+2}\})}(b_i)$. Now if $y \in V_1 \underset{\pi_{\mathcal{S}(\{a_{n+2}\})} \text{is a surjection}}{\Rightarrow} \exists x \in V$ such that $\pi_{\mathcal{S}(\{a_{n+2}\})}(x) = y$. As $\{b_i\}_{i \in \{1, \ldots, n+1\}}$ is a basis of V there exists a $\{\alpha_i\}_{i \in \{1, \ldots, n+1\}}$ with finite support such that $x = \sum_{i \in \{1, \ldots, n+1\}} \alpha_i \cdot b_i$ and thus $y = \pi_{\mathcal{S}(\{a_{n+2}\})}\left(\sum_{i \in \{1, \ldots, n+1\}} \alpha_i \cdot b_i\right) \underset{\pi_{\mathcal{S}(\{a_{n+2}\})} \text{is linear}}{=} \sum_{i \in \{1, \ldots, n+1\}} \alpha_i \cdot \pi_{\mathcal{S}(\{a_{n+2}\})}(b_i) = \sum_{i \in \{1, \ldots, n+1\}} \alpha_i \cdot \bar{b_i}$ proving that

$$\mathcal{S}(\{\bar{b_i} | i \in \{1, \ldots, n+1\}\}) = V_1 \tag{10.39}$$

For a_1 we have now the following cases to consider:

A. $(\boldsymbol{\sim_{\mathcal{S}(\{a_{n+2}\})}[a_1] = 0})$ as then neutral element in V_1 is given by $0 = \sim_{\mathcal{S}(\{a_{n+2}\})}[0]$ we have by 2.116 that $a_1 \sim_{\mathcal{S}(\{a_{n+2}\})} 0 \Rightarrow a_1 = a_1 - 0 \in \mathcal{S}(\{a_{n+2}\}) \Rightarrow \exists \gamma \in F$ such that $a_1 = \gamma \cdot a_{n+2}$ and using 10.144 we have then that A is linear dependent.

B. $(\boldsymbol{\sim_{\mathcal{S}(\{a_{n+2}\})}[a_1] \neq 0})$ from this it follows that V_1 is a non trivial vector space. Using 10.39 together with 10.156 there exists a

$$J \subseteq \{1, \ldots, n+1\} \text{ such that } \{\bar{b_i}\}_{i \in J} \text{ is a basis of } V_1 \tag{10.40}$$

As $0 \neq a_{n+2} \in V$ there exists a $\{\lambda_i\}_{i \in \{1, \ldots, n+1\}}$ such that $a_{n+2} = \sum_{i \in \{1, \ldots, n+1\}} \lambda_i \cdot b_i$ where not all $\lambda_i = 0$ [as $a_{n+2} \neq 0$]. We have then as $a_{n+2} \in \mathcal{S}(\{a_{n+2}\})$ that $0 = \sim_{\mathcal{S}(\{a_{n+2}\})}[a_{n+2}] = \pi_{\mathcal{S}(\{a_{n+2}\})}(a_{n+2}) = \pi_{\mathcal{S}(\{a_{n+2}\})}\left(\sum_{i \in \{1, \ldots, n+1\}} \lambda_i \cdot b_i\right) = \sum_{i \in \{1, \ldots, n+1\}} \lambda_i \cdot \pi_{\mathcal{S}(\{a_{n+2}\})}(b_i) = \sum_{i \in \{1, \ldots, \ldots, n+1\}} \lambda_i \cdot \bar{b_i}$ proving that $\{\bar{b_i}\}_{i \in \{1, \ldots, n+1\}}$ is linear dependent and thus that

$$J \subset \{1, \ldots, n+1\} \underset{5.46}{\Rightarrow} \#(J) < n+1 \Rightarrow \#(J) \leqslant n \tag{10.41}$$

As $\#(J) = m \leqslant n$ there exists a bijection $\sigma: \{1, \ldots, m\} \rightarrow J$ such that (see 10.149).

$$\{\bar{b}_{\sigma_i}\}_{i \in \{1, \ldots, m\}} \text{ is a basis of } V_1 \text{ where } m \leqslant n \tag{10.42}$$

If we define $\bar{A} = \{\bar{a}_i\}_{i \in \{1,\ldots,n+1\}}$ a family in V_1 where $\bar{a}_i = \pi_{\mathcal{S}(\{a_{n+2}\})}(a_i)$ then using 10.42 and $n \in S$ we have that $\bar{A} = \{\bar{a}_i\}_{i \in \{1,\ldots,n+1\}}$ is linear dependent. So there exists a $\{\varepsilon_i\}_{i \in \{1,\ldots,n+2\}}$ with not every $\varepsilon_i = 0$ such that $0 = \sum_{i \in \{1,\ldots,n+1\}} \varepsilon_i \cdot \bar{a}_i = \sum_{i \in \{1,\ldots,n+1\}} \varepsilon_i \cdot \pi_{\mathcal{S}(\{a_{n+1}\})}(a_i) = \pi_{\mathcal{S}(\{a_{n+2}\})}\left(\sum_{i \in \{1,\ldots,n+1\}} \varepsilon_i \cdot a_i\right)$ and this gives that $\sum_{i \in \{1,\ldots,n+1\}} \varepsilon_i \cdot a_i \in \mathcal{S}(\{a_{n+2}\})$. So there exists a $\mu \in F$ such that $\sum_{i \in \{1,\ldots,n+1\}} \varepsilon_i \cdot a_i = \mu \cdot a_{n+2} \Rightarrow \sum_{i \in \{1,\ldots,n+1\}} \varepsilon_i \cdot a_i - \mu \cdot a_{n+2} = 0$. if we define now $\{\kappa_i\}_{i \in \{1,\ldots,n+2\}}$ by $\kappa_i = \begin{cases} \varepsilon_i \text{ if } i \in \{1,\ldots,n+1\} \\ -\mu \text{ if } i = n+2 \end{cases}$ which is not all zeroes [as the ε_i's are]. We have then $\sum_{i \in \{1,\ldots,n+2\}} \kappa_i \cdot a_i = \sum_{i \in \{1,\ldots,n+1\}} \kappa_i \cdot a_i + \sum_{i \in \{n+2\}} \kappa_i \cdot a_i = \sum_{i \in \{1,\ldots,n+1\}} \varepsilon_i \cdot a_i - \mu \cdot a_{n+2} = 0$ proving that A is linear dependent.

As A is linear dependent in all possible cases we have then $n+1 \in S$.

Using mathematical induction we have $S = \{1, \ldots\} = \mathbb{N}$ which proves our theorem. □

Corollary 10.196. *Let V be a non trivial vector space over a field F with a basis $\{b_i\}_{i \in \{1,\ldots,n\}}$ then every independent set A is finite and $\#(A) \leqslant n$*

Proof. If A is infinite and independent then by 5.34 there exists a $C \subseteq A$ with $C \approx \mathbb{N}_0$ so there exists a bijection $c: \mathbb{N}_0 \to C$ and then $\{c_i\}_{i \in \{1,\ldots,n+1\}}$ is linear dependent by the previous axiom. Using 10.142 we have that $\{c_i | i \in \{1, \ldots, n+1\}\}$ is linear dependent. As $\{c_i | i \in \{1, \ldots, n+1\}\} \subseteq C \subseteq A$ this proves that A is linear dependent (see 10.147) contradicting the linear independence of A. So we must have that A is finite.

If now A is finite with $m = \#(A) > n$ then there exists a bijection $a: \{1, \ldots m\} \to A$, as $n < m$ we have $n + 1 \leqslant m$ and thus by the previous theorem we have that $\{a_i\}_{i \in \{1,\ldots,n+1\}}$ is linear dependent. Using 10.142 we have then that $\{a_i | i \in \{1, \ldots, n+1\}\}$ is linear dependent and thus by 10.147 we have as $\{a_i | i \in \{1, \ldots, n+1\}\} \subseteq A$ that A is linear dependent contradicting its linear independence. So we must have that $\#(A) \leqslant n$. □

Definition 10.197. *Let V be a vector space over a field F with a basis $\{b_i\}_{i \in I}$ then a basis is finite if $\{b_i | i \in I\}$ is finite. We define $\#(\{b_i\}_{i \in I}) = \#(\{b_i | i \in I\})$ and call this number the number of elements in $\{b_i\}_{i \in I}$. Note that as for a basis $\{b_i\}_{i \in I}$ we have that I and $\{b_i | i \in I\}$ are bijective (see 10.141) we have that $\{b_i | i \in I\}$ is finite if and only if I is finite and $\#(\{b_i\}_{i \in I}) = \#(I)$.*

Theorem 10.198. *Let V be a non trivial vector space over a field F with a finite basis $\{b_i\}_{i \in I}$ such that $\#\{b_i | i \in I\} = n$ then every basis $\{c_j\}_{j \in J}$ of V is finite and $\#(\{c_j\}_{j \in J}) \leqslant n$.*

Proof. If $\{b_i\}_{i\in I}$ is a finite basis with n elements then $\#(I)=n$ so there exists a bijection $h:\{1,\ldots,n\}\rightarrow I$ and using 10.149 we have that $\{b_{h(i)}\}_{i\in\{1,\ldots,n\}}$ is a basis for V. Then if $\{c_j\}_{j\in J}$ is another basis (and thus by 10.142 $\{c_j|j\in J\}$ is linear independent) we must using the previous corollary have that $\{c_j|j\in J\}$ is finite and $\#(c_j|j\in J)\leqslant n$. □

Corollary 10.199. *Let V be a non trivial vector space over a field F with a finite basis $\{b_i\}_{i\in I}$ then every other basis $\{c_j\}_{j\in J}$ is finite and $\#(\{b_i\}_{i\in I})=\#(\{c_j\}_{j\in J})$*

Proof. If $\{b_i\}_{i\in I}$ is a finite basis with n elements then by the previous theorem we have that $\{c_j\}_{j\in J}$ is also finite and $m=\#(\{c_j\}_{j\in J})\leqslant n$. Using the previous theorem again we must have as $\{c_j\}_{j\in J}$ is finite that $n=\#(\{b_i\}_{i\in I})\leqslant m$ giving that finally we have $n=m$. □

The above corollary assures us then that the following definition make sense

Definition 10.200. *Let V be a vector space over a field F then $\dim(V)$ is defined as follows*

1. *If $V=\{0\}$ (V is trivial) then $\dim(V)=0$*
2. *If V is non trivial and has a finite basis $\{b_i\}_{i\in I}$ then $\dim(V)=n$ (by the previous corollary this number is unique)*
3. *If V is non trivial and has a basis which is not finite (by the previous corollary we can then have no finite basis) then $\dim(V)=\infty$. Note that $\dim(V)=\infty$ means V is infinite and non trivial, so we assume that ∞ is different from every $n\in\mathbb{N}_0$.*

V is finite dimension if it has a finite basis or if it is trivial.

Theorem 10.201. *Let V be a vector space over F then $\dim(V)=0$ is equivalent with $V=\{0\}$*

Proof. If $V=\{0\}$ then by definition we have $\dim(V)=0$. If $\dim(V)=0$ then by definition we can not have that V is infinite dimensional so V is finite dimensional. If now $x\in V$ with $x\neq 0$ then $\{x\}$ is linear independent [if $\lambda\cdot x=0$ with $\lambda\neq 0$ then $0=\lambda^{-1}\cdot(\lambda\cdot x)=(\lambda^{-1}\cdot\lambda)\cdot x=1\cdot x=x$ contradicting $x\neq 0$]. By 10.157 there exists a B (which must be finite dimensional) such that $x\in B$ and $\{b\}_{b\in B}$ is a basis of V. Now as $x\in B\Rightarrow\#(\{b\}_{b\in B})=\#(B)>0\Rightarrow\dim(V)>0$ contradicting $\dim(V)=0$. So we must have that $x=0$ and this means that $V=\{0\}$. □

Theorem 10.202. *Let $n\in\mathbb{N}$ be $\{V_i\}_{i\in\{1,\ldots,n\}}$ be a finite family of finite dimensional vector spaces with $\forall i\in\{1,\ldots,n\}$ $\dim(V_i)=n_i$ then $\prod_{i=1}^n V_i$ (which is a vector space see 10.118) is a finite dimensional vector space with $\dim(\prod_{i=1}^n V_i)=\sum_{i=1}^n n_i$*

Proof. This follows from the construction of a basis for $\prod_{i=1}^n V_i$ (see 10.166) □

Theorem 10.203. *Let V be a vector space over F with $n=\dim(V)\in\mathbb{N}$ (hence by the above theorem a non trivial vector space) then if $\{e_i\}_{i\in\{1,\ldots,m\}}$ with $m\in\{1,\ldots,n\}$ is a linear independent set in V then there exists a basis $\{f_i\}_{i\in\{1,\ldots,n\}}$ of V such that $\forall i\in\{1,\ldots,m\}$ we have $e_i=f_i$*

Proof. If $m=n$ then the proof is trivial so we may assume that $m<n$. Take $R=\{e_i|i\in\{1,\ldots,m\}\}\subseteq V$ then using 10.143 we have that R is linear independent, using 10.154 together with $\mathcal{S}(V)=V$ we have that there exists a set B with $R\subseteq B$ such that $\{b\}_{b\in B}$ is a basis of V. As $\dim(V)=n$ we have that $\#(B)=n$ hence using 5.45 we have that $\#(K)=\#(B\setminus R)=\#(B)-\#(R)=n-m>0$. So there exists a β: $\{1,\ldots,m-n\}\to B\setminus R$ hence if we define $f:\{1,\ldots,n\}\to B$ by $f_i=f(i)=\left\{\begin{smallmatrix} e_i \text{ if } i\in\{1,\ldots,n-m\} \\ \beta((i-(n-m) \text{ if } i\in\{(n-m)+1,\ldots,n\}\end{smallmatrix}\right.$ we have $f(\{1,\ldots,n\})\subseteq R\bigcup(B\setminus R)=B$ so that $\mathcal{S}(\{f_i\}_{i\in\{1,\ldots,n\}})=\mathcal{S}(\{f_i|i\in\{1,\ldots,n\}\})=\mathcal{S}(f(\{1,\ldots,n\}))=\mathcal{S}(B)=V$ and as B is independent we have by 10.143 that $\{f_i\}_{i\in\{1,\ldots,n\}}$ is linear independent. This proves that $\{f_i\}_{i\in\{1,\ldots,n\}}$ is a basis and that $\forall i\in\{1,\ldots,m\}$ we have $e_i=f_i$. □

The above theorem and definition suggest the following definition of equality of dimensions of a vector space (taking in account the ∞).

Definition 10.204. *If V is a vector space over a field F and W a vector space over a field G then* $\dim(V)=\dim(W)$ *if and only if*

1. *V and W are both finite dimensional and* $\dim(V)=\dim(W)$
2. *V and W are both infinite dimensional*

As a consequence of this definition we have the following

Remark 10.205. If V and W are vector spaces and $\dim(V)=\dim(W)$ then we have

1. If V is trivial then W is trivial
2. If V is non trivial and finite dimensional then W is non trivial and finite dimensional with the same number of elements in its basis as V
3. If V is infinite dimensional then W is infinite dimensional

Theorem 10.206. *If V is a finite dimensional vector space over a field F with* $\dim(V)=n\in\mathbb{N}$ *then there exists a basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ of V*

Proof. As $\dim(V)\neq 0$ V is not trivial and by definition there exists a basis $\{f_i\}_{i\in I}$ with $\#(I)=n$ and as $\#(I)=n$ there exists a bijection $b:\{1,\ldots,n\}\to I$ then by 10.149 we have that $\{f_{b(i)}\}_{i\in\{1,\ldots,n\}}$ is a basis for V. So if we take $\{e_i\}_{i\in\{1,\ldots,n\}}$ where $e_i=f_{b(i)}$ we have proved our theorem. □

Theorem 10.207. *Let $n\in\mathbb{N}$, V a n-dimensional vector space over a field F and $\{e_i\}_{i\in\{1,\ldots,n\}}$ a family of independent vectors then $\{e_i\}_{i\in\{1,\ldots,n\}}$ is a basis of V*

Proof. Using 10.142 we have that $E=\{e_i|i\in\{1,\ldots,n\}\}$ is a linear independent set. Then by 10.157 we have the existence of a set B with $E\subseteq B$ and B is a basis of V so that $\mathcal{S}(B)=V$. Take now $e:\{1,\ldots,n\}\to E$ by $i\to e_i$ then we have by 10.141 that e is a bijection and thus that $\#(E)=n$. Now if $E\subset B$ then by 5.46 we have that $n=\#(E)<\#(B)\underset{\dim(V)=0}{=}n$ giving the contradiction $n<n$, so we must have $E=B$ and thus $\mathcal{S}(E)=\mathcal{S}(B)=V$ proving together with the linear independence of $\{e_i\}_{i\in\{1,\ldots,n\}}$ that $\{e_i\}_{i\in\{1,\ldots,n\}}$ is a basis of V.

□

Theorem 10.208. *Let V be a finite dimensional vector space over F and $\emptyset \neq W \subseteq V$ a sub space of V then we have that W is finite dimensional and* $\dim(W) \leqslant \dim(V)$

Proof. As $W \neq 0$ we have that W is indeed a vector space. Now we have the following cases for $\dim(V)$:

1. $\mathbf{(\dim(V) = 0)}$ then we have that $V = \{0\}$ and thus we must have that $W = \{0\}$ giving $\dim(W) = 0$ so that W is finite and $\dim(W) \leqslant \dim(V)$
2. $\mathbf{(\dim(V) \neq 0)}$ we have now the following cases for W
   a. $\mathbf{(W = \{0\})}$ then $\dim(W) = 0 \leqslant \dim(V)$ giving W is finite and $\dim(W) \leqslant \dim(V)$
   b. $\mathbf{(W \neq \{0\})}$ if now $\{b_i\}_{i \in I}$ is a basis for W then by 10.157 then there exists a B such that $\{b_i | i \in I\} \subseteq B$ and $\{b\}_{b \in B}$ is a basis for V. From the finite dimensionality of V it follows that $\{b_i | i \in I\}$ is finite so that by definition $\{b_i\}_{i \in I}$ is finite. Also using 5.45 we have that $\dim(W) = \#(\{b_i | i \in I\}) \leqslant \#(B) = \dim(V)$ so we have W is finite and $\dim(W) \leqslant \dim(V)$ □

Theorem 10.209. *Let V be a vector space over F and let $A = \mathcal{S}(\{a_1, ..., a_n\}), n \in \mathbb{N}$ then A is finite dimensional and* $\dim(A) \leqslant n$

Proof. We have to consider the following cases:

1. $\mathbf{(A = \mathcal{S}(\{0\}))}$ as $\mathcal{S}(\{0\}) = \{0\}$ we have $\dim(A) = 0 \leqslant n$
2. $\mathbf{(A = \mathcal{S}(\{a_1, ..., a_n\})}$ **with not every** $\mathbf{a_i = 0)}$ then $\mathcal{S}(\{a_1, ..., a_n\})$ is non trivial and as $\mathcal{S}(\{a_1, ..., a_n\}) = A$ there exists by 10.155 a set $B \subseteq \{a_1, ..., a_n\}$ such that $\{b\}_{b \in B}$ forms a basis of A. Using 5.45 we have that B is finite and $\dim(A) = \#(B) \leqslant n$ □

Definition 10.210. *Let V and W be vector spaces over the field F and $L: V \rightarrow W$ be a linear mapping then* $\ker(L) = \{x \in V | L(x) = 0\} = L^{-1}(\{0\})$. *We call* $\ker(L)$ *the kernel of L. Note that as $L(0) = 0$ we have* $0 \in \ker(L)$

Theorem 10.211. *Let V and W be vector spaces over the field F and let $L: V \rightarrow W$ be a linear mapping then L is regular (see 10.172) if and only if* $\ker(L) = \{0\}$

Proof. if $\ker(L = \{0\})$ then if $x, y \in W$ such that $L(x) = L(y) \Rightarrow 0 = L(x) - L(y) = L(x - y) \Rightarrow x - y \in \ker(L) = \{0\} \Rightarrow x - y = 0 \Rightarrow x = y$, proving the regularity of L. If L is injective and $x \in \ker(L)$ then $0 = L(x) \underset{L(0)=0}{\Rightarrow} L(0) = L(x) \Rightarrow x = 0$ if L is injective so that $\ker(L) = \{0\}$ □

Theorem 10.212. *Let V, W be vector spaces over a field F with* $\dim(V) = n \in \mathbb{N}$ *(so V is finite dimensional and non trivial), $L: V \rightarrow W$ a linear mapping and $\{b_i\}_{i \in I}$ a basis for V then we have that L is regular if and only if $\{L(b_i)\}_{i \in I}$ is linear independent. Furthermore if* $\dim(W) = n$ *then we have that L is regular if and only if L is a isomorphism.*

Proof.

1. If L is regular (=injective) then if $\{\alpha_i\}_{i\in L}$ is a family with finite support such that $0=\sum_{i\in I}\alpha_i\cdot L(b_i)=\sum_{i\in I}L(\alpha_i\cdot b_i)\underset{L\text{ is linear and }10.174}{=}L(\sum_{i\in I}\alpha_i\cdot b_i)\Rightarrow\sum_{i\in I}\alpha_i\cdot b_i\in\ker(L)\underset{L\text{ is regular and previous theorem}}{=}\{0\}\Rightarrow\sum_{i\in I}\alpha_i\cdot b_i\underset{\{b_i\}_{i\in I}\text{ is a basis}}{\Rightarrow}\forall i\in I$ we have $\alpha_i=0$ proving that $\{L(b_i)\}_{i\in I}$ is linear independent.
2. If $\{L(b_i)\}_{i\in I}$ is linear independent then if $x\in\ker(L)$ we have $L(x)=0$, as $x\in V$ and $\{b_i\}_{i\in I}$ is a basis for V there exists a $\{\alpha_i\}_{i\in I}$ with finite support such that $x=\sum_{i\in I}\alpha_i\cdot b_i\Rightarrow 0=L(x)=L(\sum_{i\in I}\alpha_i\cdot b_i)=\sum_{i\in I}\alpha_i\cdot L(b_i)\underset{\{L(b_i)\}_{i\in I}\text{ is linear independent}}{\Rightarrow}\forall i\in I$ we have $\alpha_i=0\Rightarrow x=\sum_{i\in I}\alpha_i\cdot b_i=\sum_{i\in I}0\cdot b_i=0$ so that $\ker(L)=\{0\}$ proving using the above theorem that L is regular.
3. If L is injective (=regular) then $L:\{b_i|i\in I|\}\rightarrow\{L(b_i)|i\in L\}=L(\{b_i|i\in I\})$ is a bijection and thus $\dim(V)=n=\#(\{b_i|i\in I\})=\#(\{L(b_i)|i\in I\})$. By (1) we have that $\{L(b_i)\}_{i\in I}$ is linear independent and using 10.157 there exists a B' such that $\{L(b_i)|i\in I\}\subseteq B'$ and $\{b\}_{b\in B'}$ is a basis for W and thus $\#(B')=\dim(W)=n$. If now $\{L(b_i)|i\in I|\}\subset B'$ then by 5.46 we have $n=\#(\{L(b_i)|i\in I\})<\#(B')=n$ a contradiction. So we have $\{L(b_i)|i\in I\}=B'$ giving that $\mathcal{S}(\{L(b_i)|i\in I\})=\mathcal{S}(B')=W$ and this together with the linear independence of $\{L(b_i)\}_{i\in I}$ proves that $\{L(b_i)_{i\in I}\}$ is a basis of W. So if $y\in W$ then there exists a family $\{\alpha_i\}_{i\in I}$ with finite support such that $y=\sum_{i\in I}\alpha_i\cdot L(b_i)=L(\sum_{i\in I}\alpha_i\cdot b_i)\in L(V)$ proving that $W\subseteq L(V)$ and thus that L is surjective. Together with injectivity and linearity this means that L is a isomorphism.
4. If L is a isomorphism then evidently it is also injective. □

Theorem 10.213. *Let V and W be vector spaces over a field F and $L:V\rightarrow W$ a linear mapping then $\ker(L)$ is a subspace of V. As $0\in\ker(L)$ this also implies that $\ker(L)$ is a vector space.*

Proof. If $x,y\in\ker(L)$ and $\alpha,\beta\in F$ then $L(\alpha\cdot x+\beta\cdot y)=\alpha\cdot L(x)+\beta\cdot L(y)=\alpha\cdot 0+\beta\cdot 0=0$ so $\alpha\cdot x+\beta\cdot y\in\ker(L)$ □

Theorem 10.214. *Let V and W be vector spaces over a field F and $L:V\rightarrow W$ a linear mapping then $L(V)$ is a subspace of W. As $0=L(0)\in L(V)$ this also implies that $L(V)$ is a vector space.*

Proof. If $x,y\in L(V)$ and $\alpha,\beta\in F$ then $\exists x',y'\in V$ such that $x=L(x'),y=L(y')$ then $L(V)\ni L(\alpha\cdot x'+\beta\cdot y')=\alpha\cdot L(x')+\beta\cdot L(y')=\alpha\cdot x+\beta\cdot y$ □

Definition 10.215. *Let V,W be vector spaces over a field F then* $\operatorname{rank}(L)=\dim(L(V))$

Theorem 10.216. *Let V,W be vector spaces over a field F and $L:V\rightarrow W$ a linear mapping then if V is finite dimensional we have that $L(V)$ is finite dimensional and $\dim(L(V))=\operatorname{rank}(L)\leqslant\dim(V)$. Furthermore if W is finite dimensional then $\operatorname{rank}(L)\leqslant\dim(W)$*

Proof. First we consider the trivial case where $\dim(V)=0$, then $V=\{0\}$ and thus $L(V)=\{0\} \Rightarrow \operatorname{rank}(L)=\dim(L(V))=0 \leqslant \dim(V)=0$, also if W is finite dimensional then $\operatorname{rank}(L)=0 \leqslant \dim(W)$. Second take the non trivial case and assume that $\{b_i\}_{i\in I}$ is a basis of V with $\#(\{b_i|i\in I\})=n$ then if $y\in L(V)$ there exists a $x\in V$ such that $y=L(x)$ and a $\{\alpha_i\}_{i\in I}$ with finite support such that $x=\sum_{i\in I}\alpha_i\cdot b_i$ so that $y=L(x)=L(\sum_{i\in I}\alpha_i\cdot b_i)=\sum_{i\in I}\alpha_i\cdot L(b_i)$ proving that

$$y\in L(V) \text{ then } \exists\{\alpha_i\}_{i\in I} \text{ with finite support such that } y=\sum_{i\in I}\alpha_i\cdot L(b_i) \tag{10.43}$$

Now given the fact that $L:\{b_i|i\in I\}\rightarrow L(\{b_i|i\in I)=\{L(b_i)|i\in I\}$ is a surjection and the finiteness of $\{b_i|i\in I\}$ we have by 5.48 that

$$L(\{b_i|i\in I\}) \text{ is finite and } \#(\{L(b_i)|i\in I\}) \leqslant \#(\{b_i|i\in I\})=\dim(V) \tag{10.44}$$

So let $m=\#(\{L(b_i)|i\in I\})$ then $m\leqslant\dim(V)$ and there exists a bijection $h:\{0,\ldots,m-1\}\rightarrow\{L(b_i)|i\in I\}$. Define then the family $\{I_i\}_{i\in\{0,\ldots,m-1\}}$ where $I_i=\{j\in I|L(b_j)=h(i)\}$. If now $i,j\in\{0,\ldots,m-1\}$ with $i\neq j$ then if $k\in I_i\bigcap I_j$ we have $L(b_k)=h(i)$, $h(j)\Rightarrow h(i)=h(j)\Rightarrow i=j$ a contradiction so we must have that $I_i\bigcap I_j=\emptyset$. Also as $I_i\subseteq I$ we have $\bigcup_{i\in\{0,\ldots,m-1\}}I_i\subseteq I$. If $k\in I$ then $L(b_k)\in\{L(b_i)|i\in I\}\Rightarrow h^{-1}(L(b_k))\in\{0,\ldots,m-1\}$ and $h(h^{-1}(L(b_k)))=L(b_k)\Rightarrow k\in I_{h^{-1}(L(b_k))}\in\bigcup_{i\in\{0,\ldots,m-1\}}I_i$ proving that $I\subseteq\bigcup_{i\in\{0,\ldots,m-1\}}I_i$ and finally that $I=\bigcup_{i\in\{0,\ldots,m-1\}}I_i$ where I_i are disjoint sets. If now $y\in L(V)$ then by 10.43 there exists a $\{\alpha_i\}_{i\in I}$ such that $y=\sum_{i\in I}\alpha_i\cdot L(b_i)\underset{10.46}{=}\sum_{i=0}^{m-1}(\sum_{k\in I_i}\alpha_k\cdot L(b_k))=\sum_{i=0}^{m-1}(\sum_{k\in I_i}\alpha_k\cdot h(i))=\sum_{i=0}^{m-1}(\sum_{k\in I_i}\alpha_k)\cdot h(i)=\sum_{i=0}^{m-1}\beta_{h(i)}\cdot h(i)=\sum_{v\in\{L(b_i)|i\in I\}}\beta_v\cdot v$ if we define $\{\beta_v\}_{v\in\{L(b_i|i\in I)\}}$ by $\beta_v=\left(\sum_{k\in I_{h^{-1}(v)}}\alpha_k\right)$ so that $\beta_{h(i)}=\left(\sum_{k\in I_{h^{-1}(h(i))}}\alpha_k\right)=\sum_{k\in I_i}\alpha_k$. This proves that $y\in\mathcal{S}(\{L(b_i|i\in I)\})$ and thus

$$L(V)\subseteq\mathcal{S}(\{L(b_i)|i\in I\}) \tag{10.45}$$

As $L(\{b_i|i\in I\})\subseteq V$ we have by 10.136 that $\mathcal{S}(\{b_i|i\in I\})\subseteq\mathcal{S}(L(V))\underset{L(V)\text{ is a vector space}}{=}L(V)$ and this gives with 10.45 that

$$L(V)=\mathcal{S}(\{L(b_i)|i\in I\}) \tag{10.46}$$

Using 10.155 there exists a $E\subseteq\{L(b_i)|i\in I\}$ such that $\{e\}_{e\in E}$ is a basis of $L(V)$ using 5.45 and 10.44 we have that E is finite and $\dim(L(V))=\#(E)\leqslant\#(\{L(b_i)|i\in I\})\leqslant\dim(V)$. This proves that $L(V)$ is finite dimensional and that

$$\dim(L(V))=\operatorname{rank}(L)\leqslant\dim(V) \tag{10.47}$$

Lastly if W is finite dimensional and as $L(V)$ is a subspace of W we have by 10.208 that

$$\operatorname{rank}(L)=\dim(L(V))\leqslant\dim(W)$$

□

Theorem 10.217. *Let V,W be vector spaces over a field F, V is finite dimensional and $L:V\rightarrow W$ a linear mapping then L is surjective if and only if W is finite dimensional and* $\operatorname{rank}(L)=\dim(W)$

Proof. If L is surjective then $L(V) = W$ and by using the previous theorem we have that $L(V)$ and thus W is finite dimensional and $\operatorname{rank}(L) = \dim(L(V)) = \dim(W)$.

If W is finite dimensional and $\operatorname{rank}(L) = \dim(W)$ then $\dim(L(V)) = \dim(W)$, we must then consider the following two cases:

1. **($\dim(W) = 0$)** then $W = \{0\}$ and we must have $L(V) = \{0\}$ so $L: \{0\} \to \{0\}$ is as $L(0) = 0$ surjective.
2. **($\dim(W) \neq 0$)** then $0 \neq \operatorname{rank}(L) = \dim(L(V))$ so there exists a basis $\{b_i\}_{i \in I}$ of $L(V)$ such that $\#(\{b_i | i \in I\}) = \dim(W)$. Using 10.157 there exists a B such that $\{b_i | i \in I\} \subseteq B$ and $\{b\}_{b \in B}$ is a basis of W. If now $\{b_i | i \in I\} \subset B$ then by 5.46 we would have the contradiction $\dim(W) = \#(\{b_i | i \in I\}) < \#(B) = \operatorname{dimw}(W)$ so we must have that $\{b_i | i \in I\} = B$. So $L(V) = \mathcal{S}(\{b_i | i \in I\}) = \mathcal{S}(B) = W$ proving surjectivity. □

Theorem 10.218. *Let V, W be vector spaces over a field F, V is finite dimensional and $L: V \to W$ a linear mapping, then* $\dim(V) = \dim(\ker(L)) + \operatorname{rank}(L)$

Proof. We divide the proof in two cases:

1. **($\ker(L) = \{0\}$)** then by 10.211 we have that L is injective, also we have that $\dim(\ker(L)) = 0$. Let now $\{b_i\}_{i \in I}$ be a basis of V then we have $\dim(V) = \#(\{b_i | i \in I\})$. Using 10.212 we have then by regularity of the basis that $\{L(b_i)\}_{i \in I}$ is linear independent. If $y \in L(V)$ then there exists a $x \in V$ such that $y = L(x)$ and a $\{\alpha_i\}_{i \in I}$ with finite support such that $x = \sum_{i \in I} \alpha_i \cdot b_i$ and thus we have $y = L(\sum_{i \in I} \alpha_i \cdot b_i) = \sum_{i \in I} \alpha_i \cdot L(b_i)$ proving that $L(V) \subseteq \mathcal{S}(\{L(b_i) | i \in I\}) \subseteq L(V) \Rightarrow L(V) = \mathcal{S}(\{L(b_i) | i \in I\})$. This proves that $\{L(b_i)\}_{i \in I}$ is a basis for $L(V)$ and thus $\operatorname{rank}(L) = \dim(L(V)) = \#(\{b_i | i \in I\}) = \dim(V) = \dim(V)$. As $\dim(\ker(L)) = 0$ we have then finally $\dim(\ker(L)) + \operatorname{rank}(L) = \dim(V)$.
2. **($\ker(L) \neq \{0\}$)** Let $\{b_i\}_{i \in I}$ be a basis for $\ker(L)$, define then $B = \{b_i | i \in I\}$, and using 10.157 extend it to a B' such that $B \subseteq B'$ and $\{w\}_{w \in B'}$ is a basis for V. As V is finite dimensional we have that B' and thus also B and $B' \setminus B$ is finite, furthermore we have (see 5.45) that

$$\#(B' \setminus B) + \#(B) = \#(B') \tag{10.48}$$

   We prove now that

$$L_{|B' \setminus B}: B' \setminus B \to L(B' \setminus B) \text{ is } a \text{ bijection} \tag{10.49}$$

   Proof.

   a. **(injectivity)** If $x, y \in B' \setminus B$ is such that $L_{|B' \setminus B}(x) = L_{|B' \setminus B}(y) \Rightarrow L(x) = L(y) \Rightarrow L(x) - L(y) = 0 \Rightarrow L(x - y) = 0 \Rightarrow x - y \in \ker(L)$. As B is a basis for $\ker(L)$ (see 10.142 and $\{b_i\}_{i \in I}$ is a basis of $\ker(L)$). There exists a family $\{\alpha_w\}_{w \in B}$ such that $x - y = \sum_{w \in B} \alpha_w \cdot w$ and thus

$$0 = \sum_{w \in B} \alpha_w \cdot w + 1 \cdot y + (-1) \cdot x \tag{10.50}$$

Assume now that $x \neq 0$ and define using the fact that $x, y \notin B$ the family $\{\lambda_w\}_{w \in B'}$ by $\lambda_w = \begin{cases} \alpha_w \text{ if } w \in B \\ 1 \text{ if } w \in \{y\} \\ -1 \text{ if } w \in \{x\} \\ 0 \text{ if } w \in B' \setminus (B \bigcup \{x, y\}) \end{cases}$ then we have $\sum_{w \in B'} \lambda_w \cdot w = \sum_{w \in B} \lambda_w \cdot w + \sum_{w \in \{y\}} \lambda_w \cdot w + \sum_{w \in \{x\}} \lambda_w \cdot w + \sum_{w \in B' \setminus (B \bigcup \{x, y\})} \lambda_w \cdot w = \sum_{w \in B} \alpha_w \cdot w + 1 \cdot y + (-1) \cdot x + 0 \underset{10.50}{=} 0$. Using the fact that B' is a basis we must have that $\forall w \in W \vDash \lambda_w = 0 \Rightarrow \forall w \in B$ we have $\alpha_w = \lambda_w = 0$ proving that $x - y = 0 \Rightarrow x = y$ contradicting the assumption that $x \neq y$. So we must have that $x = y$ proving that $L_{|B' \setminus B}$ is bijective.

b. **(surjectivity)** This is trivial □

Now if $y \in L(V)$ there exists a $x \in V$ and $\{\alpha_w\}_{w \in B'}$ such that $y = L(x)$ and $x = \sum_{w \in B'} \alpha_w \cdot w$ so that $y = L(\sum_{w \in B'} \alpha_w \cdot w) = \sum_{w \in B'} \alpha_w \cdot L(w) = \sum_{w \in B' \setminus B} \alpha_w \cdot L(w) + \sum_{w \in B} \alpha_w \cdot L(w) = \sum_{w \in B' \setminus B} \alpha_w \cdot w + \sum_{w \in B} \alpha_w \cdot 0 = \sum_{w \in B' \setminus B} \alpha_w \cdot L(w) = \sum_{w \in B' \setminus B} \alpha'_{L(w)} \cdot L(w) = \sum_{w \in L(B' \setminus B)} \alpha'_w \cdot w$ [where $\{\alpha'_w\}_{w \in L(B' \setminus B)}$ is defined by $\alpha'_w = \alpha_{L^{-1}_{|B' \setminus B}(w)}$ so that $\alpha'_{L(w)} = \alpha_{L^{-1}_{|B' \setminus B}(L(w))} = \alpha_w$]. This proves that

$$L(V) \subseteq \mathcal{S}(L(B' \setminus B)) \tag{10.51}$$

As we have also $B' \setminus B \subseteq V \Rightarrow L(B' \setminus B) \subseteq L(V) \Rightarrow \mathcal{S}(L(B' \setminus B)) \subseteq \mathcal{S}(L(V)) \underset{L(V) \text{ is a vector space}}{=} L(V)$ and this together with 10.51 gives

$$L(V) = \mathcal{S}(L(B' \setminus B)) \tag{10.52}$$

Now if $\{\lambda_w\}_{w \in L(B' \setminus B)}$ is a family in F with finite support such that $0 = \sum_{w \in L(B' \setminus B)} \lambda_w \cdot w \underset{L_{|B' \setminus B}: B' \setminus B \rightarrow L(B' \setminus B) \text{ is a bijection}}{=} \sum_{w \in B' \setminus B} \lambda_{L_{|B' \setminus B}(w)} \cdot L_{|B' \setminus B}(w) = \sum_{w \in B' \setminus B} \lambda_{L(w)} \cdot L(w) = L\left(\sum_{w \in B' \setminus B} \lambda_{L(w)} \cdot w\right)$ so $\sum_{w \in B' \setminus B} \lambda_{L(w)} \cdot w \in \ker(L)$, so there exists a $\{\beta_w\}_{w \in B}$ such that $\sum_{w \in B' \setminus B} \lambda_{L(w)} \cdot w = \sum_{w \in B} \beta_w \cdot w \Rightarrow \sum_{w \in B' \setminus B} \lambda_{L(w)} \cdot w + \sum_{w \in B} (-\beta_w) \cdot w = 0$. Define now $\{\gamma_w\}_{w \in B'}$ by $\gamma_w = \begin{cases} \lambda_{L(w)} \text{ if } w \in B' \setminus B \\ -\beta_w \text{ if } w \in B \end{cases}$ then $\sum_{w \in B'} \gamma_w \cdot w = \sum_{w \in B' \setminus B} \gamma_w \cdot w + \sum_{w \in B} \gamma_w \cdot w = \sum_{w \in B' \setminus B} \lambda_{L(w)} \cdot w + \sum_{w \in B} (-\beta_w) \cdot w = 0$ and by the independence of B' we must have $\forall w \in B'$ that $\gamma_w = 0 \Rightarrow \forall w \in B' \setminus B$ we have $\lambda_{L(w)} = 0$. If now $v \in L(B' \setminus B)$ then $\exists v' \in B' \setminus B$ such that $v = L(v')$ and $\lambda_v = \lambda_{L(v')} = 0$ proving that

$$L(B' \setminus B) \text{ is linear independent} \tag{10.53}$$

Using 10.52 and 10.53 we have that $L(B' \setminus B)$ is a basis for $L(V)$ and thus we have $\dim(L(V)) = \#(L(B' \setminus B)) \underset{L_{|B' \setminus B}: B' \setminus B \rightarrow L(B' \setminus B) \text{ is a bijection}}{=} \#(B' \setminus B)$ and thus $\dim(L(V)) + \dim(\ker(L)) = \#(B' \setminus B) + \#(B) \underset{10.48}{=} \#(B') = \dim(V)$ so we have finally

$$\dim(L(V)) + \dim(\ker(L)) = \dim(V)$$ □

or

$$\operatorname{rank}(L)+\dim(\ker(L))=\dim(V) \tag{10.54}$$

Corollary 10.219. *Let V, W be finite dimensional vector spaces over a field F and $L: V\rightarrow W$ a regular linear mapping then* $\dim(V)\leqslant\dim(W)$

Proof. $\dim(V) = \dim(\ker(L)) + \operatorname{rank}(L) \underset{\ker(L)=\{0\}}{=} 0 + \operatorname{rank}(L) = \operatorname{rank}(L) \leqslant \dim(W)$ □

Corollary 10.220. *Let V, W be vector spaces over a field F, V is finite dimensional and $L\in\operatorname{Hom}(V,W)$ then L is injective if and only if* $\operatorname{rank}(L)=\dim(V)$*. Furthermore L is a isomorphism if and only if* $\dim(W)=\dim(V)$ *and* $\operatorname{rank}(L)=\dim(V)$

Proof.

1. Assume that L is injective then $\ker(L)=0\Rightarrow\dim(V)=\operatorname{rank}(L)$
2. Assume that $\operatorname{rank}(L)=\dim(V)$ then by 10.218 we have that $\ker(L)=0$ which by 10.201 means that $\ker(L)=\{0\}$ proving that L is regular.
3. Assume that L is a isomorphism then L is surjective and by 10.217 we have that W is finite dimensional and $\operatorname{rank}(L)=\dim(W)$. From injectivity and (1) we have that $\dim(V)=\operatorname{rank}(L)$ and thus $\dim(W)=\dim(V)$.
4. Assume that $\dim(W)=\dim(V)$ and $\operatorname{rank}(L)=\dim(V)$. Form $\dim(V)=\dim(W)$ and the fact that V is finite dimensional we must have that W is finite dimensional. From $\operatorname{rank}(L)=\dim(V)=\dim(W)$ we have by 10.217 that L is surjective. By (2) we have also that L is injective and thus L is bijective and thus a isomorphism. □

Theorem 10.221. *Let U, V and W be vector spaces over a field F where V is finite dimensional then we have:*

1. *If $L: U\rightarrow V$ is a linear mapping and $H: V\rightarrow W$ is a regular linear mapping then* $\operatorname{rank}(L)=\operatorname{rank}(H\circ L)$
2. *If $L: U\rightarrow V$ is a isomorphism and $H: V\rightarrow W$ is a linear mapping then* $\operatorname{rank}(H)=\operatorname{rank}(H\circ L)$

Proof.

1. If $H: V\rightarrow W$ is injective and linear this means that $H_{|L(U)}: L(U)\rightarrow H(L(U))$ is a isomorphism (by 10.214 $L(U)$ is a vector space). Using 10.220 and the fact that $L(U)$ is finite dimensional (see 10.208 and the finite dimensionality of V) we have then that $\dim(L(U))=\dim(H(L(U)))$ so that $\operatorname{rank}(H\circ L)=\dim(H(L(U)))=\dim(L(U))=\operatorname{rank}(L)$
2. If $x\in\ker(H\circ L)\Rightarrow H(L(x))=0\Rightarrow L(x)\in\ker(H)\Rightarrow x\in L^{-1}(\ker(H))$ then
$$\ker(H\circ L)\subseteq L^{-1}(\ker(H)) \tag{10.55}$$
Also if $x\in L^{-1}(\ker(H))\Rightarrow L(x)\in\ker(H)\Rightarrow H(L(x))=0\Rightarrow x\in\ker(H\circ L)$ this gives together with 10.55 that
$$\ker(H\circ L)=L^{-1}(\ker(H)) \tag{10.56}$$

From the fact that L is a bijection we have that $L_{|L^{-1}(\ker(H))}: L^{-1}(\ker(H)) = \ker(H \circ L) \rightarrow L(L^{-1}(\ker(H))) = \ker(H)$ is a bijection and as $\ker(H \circ L)$ and $\ker(K)$ are vector spaces we have that $L_{|\ker(H \circ L)} = L_{|L^{-1}(\ker(H))}: \ker(H \circ L) \rightarrow \ker(H)$ is a isomorphism. Using 10.220 we have then that

$$\dim(\ker(H \circ L)) = \dim(\ker(H)) \tag{10.57}$$

Now as by 10.181 $L^{-1}: V \rightarrow U$ is a isomorphism and using 10.220 and the finite dimensionality of V we have

$$\dim(V) = \dim(U) \text{ (so } U \text{ is finite)} \tag{10.58}$$

Now as V is finite dimensional we can use 10.218 to get

$$\dim(\ker(H)) + \operatorname{rank}(H) = \dim(V) \tag{10.59}$$

From the finiteness of U(see 10.59) and the linearity (see 10.182) of $H \circ L$: $U \rightarrow W$ and again 10.218 we have

$$\dim(\ker(H \circ L)) + \operatorname{rank}(H \circ L) = \dim(U) \tag{10.60}$$

So using 10.60, 10.57 we have $\dim(\ker(H)) + \operatorname{rank}(H \circ L) = \dim(U) \underset{10.59}{=} \dim(V) \underset{10.59}{=} \dim(\ker(H)) + \operatorname{rank}(H)$ giving that

$$\operatorname{rank}(H \circ L) = \operatorname{rank}(H)$$ □

Theorem 10.222. *Given $n, m \in \mathbb{N}_0$, F a field and F^n a vector spaces over F (defined in 10.116), Y a vector space over F and $\sigma \in P_n$ (see 10.65 then if $L: F^n \rightarrow Y$ is a linear mapping then we have the following:*

1. *σL is a linear mapping (see 10.82)*
2. *If L is a isomorphism then σL is a isomorphism*

Proof.

1. **(linearity)** If $x = (x_1, \ldots, x_n) \in F^n, y = (y_1, \ldots, y_n) \in F^n$ [so actually $x: \{1, \ldots, n\} \rightarrow F$, $y: \{1, \ldots, n\} \rightarrow F$] and $\alpha, \beta \in F$ then $\forall i \in \{1, \ldots, n\}$ we have $((\alpha \cdot x + \beta \cdot y) \circ \sigma)(i) = (\alpha \cdot x + \beta \cdot y)(\sigma(i)) = \alpha \cdot x(\sigma(i)) + \beta \cdot y(\sigma(i)) = \alpha \cdot (x \circ \sigma)(i) + \beta \cdot (y \circ \sigma)(i) = (\alpha \cdot (x \circ \sigma) + \beta \cdot (y \circ \sigma))(i)$ so that $(\alpha \cdot x + \beta \cdot y) \circ \sigma = \alpha \cdot (x \circ \sigma) + \beta \cdot (y \circ \sigma)$. So $\sigma L(\alpha \cdot x + \beta \cdot y) = L((\alpha \cdot x + \beta \cdot y) \circ \sigma) = L(\alpha \cdot (x \circ \sigma) + \beta \cdot (y \circ \sigma)) = \alpha \cdot L(x \circ \sigma) + \beta \cdot L(y \circ \sigma) = \alpha \cdot \sigma L(x) + \beta \cdot \sigma L(y)$
2. **(bijectivity)**
   a. **(injectivity)** Let $x, y \in F^n$ is such that $\sigma L(x) = \sigma L(y) \Rightarrow L(x \circ \sigma) = L(y \circ \sigma) \underset{L \text{ is isomorphism}}{\Rightarrow} x \circ \sigma = y \circ \sigma \Rightarrow (x \circ \sigma) \circ \sigma^{-1} = (y \circ \sigma) \circ \sigma^{-1} \Rightarrow x \circ (\sigma \circ \sigma^{-1}) = y \circ (\sigma \circ \sigma^{-1}) \Rightarrow x = y$
   b. **(surjectivity)** Let $y \in Y$ then by surjectivity of L there exists a $x: \{1, \ldots, n\} \rightarrow F$ such that $L(x) = y$. Taken then $x \circ \sigma^{-1}: \{1, \ldots, n\} \rightarrow F$ so that $\sigma L(x \circ \sigma^{-1}) = L((x \circ \sigma^{-1}) \circ \sigma) = L(x \circ (\sigma^{-1} \circ \sigma)) = L(x) = y$ □

Theorem 10.223. **(Universal property)** *Let X be a vector space over the field F, $\emptyset \neq Y \subseteq X$ a sub vector space of X and Z another vector space over the field F and $\varphi: X \rightarrow Z$ a linear mapping such that $Y \subseteq \ker(\varphi)$ then there exists a* ***unique*** *linear mapping $h: X/Y \rightarrow Z$ such that $\varphi = h \circ \pi_Y$*

Proof. Define $h: X/Y \rightarrow Z$ by $h(\sim_Y[x]) \rightarrow \varphi(x)$, this is a well defined function for if $\sim_Y[x] = \sim_Y[x']$ then $x - x' \in Y \subseteq \ker(\varphi) \Rightarrow \varphi(x - x') = 0 \underset{\varphi \text{ is linear}}{\Rightarrow} \varphi(x) - \varphi(x') = 0 \Rightarrow \varphi(x) = \varphi(x')$ so that $h(\sim_Y[x]) = \varphi(x) = \varphi(x') = h(\sim_Y[x'])$. Also h is linear as $h(\alpha \cdot \sim_Y[x] + \beta \cdot \sim_Y[y]) = h(\sim_Y[\alpha \cdot x + \beta \cdot y]) = \varphi(\alpha \cdot x + \beta \cdot y) \underset{\varphi \text{ is linear}}{=} \alpha \cdot \varphi(x) + \beta \cdot \varphi(y) = \alpha \cdot h(\sim_Y[x]) + \beta \cdot h(\sim_Y[y])$. Next if $x \in X$ then $\pi_Y[x] = \sim_Y[x]$ so that $(h \circ \pi_Y)(x) = h(\pi_Y(x)) = h(\sim_Y[x]) = \varphi(x)$ so that $h \circ \pi_Y = \varphi$. Finally if $h': X/Y \rightarrow Z$ is such that $h' \circ \pi_Y = \varphi$ then $\forall \sim_Y[x] \in X/Y$ we have $h(\sim_Y[x]) = \varphi(x) = h'(\pi_Y(x)) = h'(\sim_Y[x])$ so that $h = h'$ proving uniqueness. □

10.7 Multilinear Mappings

Definition 10.224. *Let $n \in \mathbb{N}$ and $i \in \{1, ..., n\}$ then we define $\{1, ..., i-1, i+1, ..., n\}$ to be $\{1, ..., n\} \setminus \{i\}$. If $i, j \in \{1, ..., n\}$ then we define $\{1, ..., i-1, i+1, ..., j-1, j+1, ...n\}$ by $\{1, ..., n\} \setminus \{i, j\}$.*

Definition 10.225. *If $n \in \mathbb{N}$ and $\{X_i\}_{i \in \{1, ..., n\}}$ a family of sets and $x = (x_1, ..., x_n) \in \prod_{i \in \{1, ..., n\}} X_i$ is such that $x_i = a$ for $i \in \{1, ..., n\}$ then we note this as $(x_1, ..., x_{i-1}, a, x_{i+1}, ..., x_n)$. In the special cases that $i = 1$ or $i = n$ we sometimes use $(a, x_2, ..., x_n)$ or $(x_1, ..., x_{n-1}, a)$ instead of $(x_1, ..., x_{i-1}, a, x_{i+1}, ...x_n)$.*

Definition 10.226. *If $n \in \mathbb{N}$ an $\{X_i\}_{i \in \{1, ..., n\}}$ a family of sets then if $i \in \{1, ..., n\}$, $a \in X_i$ and $x \in \prod_{i \in \{1, ..., n\}}$ then we define $x_{i \rightarrow a}$ by $(x_{i \rightarrow a})_j = \begin{cases} x_i \text{ if } j \in \{1, ..., n\} \setminus \{i\} \\ a \text{ if } j = i \end{cases}$ (or using the above definition we have $(x_1, ..., x_n)_{i \rightarrow a} = (x_1, ..., x_{i-1}, a, x_{i+1}, ..., x_n)$*

Theorem 10.227. *If $n \in \mathbb{N}$ an $\{X_i\}_{i \in \{1, ..., n\}}$ a family of sets then if $i \in \{1, ..., n\}$, $a \in X_i$ and $x \in \prod_{i \in \{1, ..., n\}} X_i$ then if $i, j \in \{1, ..., n\}$ and $a \in X_i, b \in X_j$ we have $(x_{i \rightarrow a})_{j \rightarrow b} = (x_{j \rightarrow b})_{i \rightarrow a}$*

Proof. If $k \in \{1, ..., n\}$ then we have $[(x_{i \rightarrow a})_{j \rightarrow b}]_k = \begin{cases} (x_{i \rightarrow a})_k \text{ if } k \in \{1, ..., n\} \setminus \{j\} \\ b \text{ if } k = j \end{cases} = \begin{cases} x_k \text{ if } k \in \{1, ..., n\} \setminus \{j, i\} \\ a \text{ if } k = i \\ b \text{ if } k = j \end{cases} = \begin{cases} (x_{j \rightarrow b})_k \text{ if } k \in \{1, ..., n\} \setminus \{i\} \\ a \text{ if } k = i \end{cases} = [(x_{j \rightarrow b})_{i \rightarrow a}]_k$ so we have $(x_{i \rightarrow a})_{j \rightarrow b} = (x_{j \rightarrow b})_{i \rightarrow a}$ □

Definition 10.228. *Let $n \in \mathbb{N}$ and $\{X_i\}_{i \in \{1, ..., n\}}$ a family of vector spaces over a field F, $\prod_{i \in \{1, ..., n\}} X_i$ the product vector space over the field F and Y a vector space over the field F then a mapping $L: \prod_{i \in \{1, ..., n\}} X_i \rightarrow Y$ is* ***multilinear*** *if $\forall i \in \{1, ..., n\}$ and $\forall x \in \prod_{i \in \{1, ..., n\}} X_i$ with $x_i = \alpha \cdot y + \beta \cdot z$ where $y, z \in X_i$ and $\alpha, \beta \in F$ we have $L(x) = \alpha \cdot L(x_{i \rightarrow y}) + \beta \cdot L(x_{i \rightarrow z})$. Or using the definitions above we say that $L: \prod_{i \in \{1, ..., n\}} X_i \rightarrow Y$ is multilinear if $\forall i \in \{1, ..., n\}$ and $\forall (x_1, ..., x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, ..., x_n) \in \prod_{i \in \{1, ..., n\}} X_i$ where $x, y \in X_i$ and $\alpha, \beta \in F$ we have $L(x_1, ..., x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, ..., x_n) = \alpha \cdot L(x_1, ..., x_{i-1}, x, x_{i+1}, ..., x_n) + \beta \cdot L(x_1, ..., x_{i-1}, y, x_{i+1}, ..., x_n)$*

Example 10.229. $0: \prod_{i \in \{1, \ldots, n\}} X_i \rightarrow Y$ is multilinear

Proof. If $(x_1, \ldots, x_{i-1}, \alpha \cdot y + \beta \cdot z, x_{i+1}, \ldots, x_n)$ then $0(x_1, \ldots, x_{i-1}, \alpha \cdot y + \beta \cdot z, x_{i+1}, \ldots, x_n) = 0 = 0 + 0 = \alpha \cdot 0 + \beta \cdot 0 = \alpha \cdot 0\,(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) + \beta \cdot L(x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_n)$ □

Example 10.230. Let $n \in \mathbb{N}$ and F a field then $\odot: F^n \rightarrow F$ defined by $x = (x_1, \ldots, x_n) \rightarrow \odot(x_1, \ldots, x_n) = \prod_{i \in \{1, \ldots, n\}} x_i$ is a multilinear mapping.

Proof. We prove this by induction on n so let $S = \left\{n \in \mathbb{N} | \text{If } (x_1, \ldots, x_i, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_n) \in F^n \text{ then } \prod_{j \in \{1, \ldots, n\}} (x_1, \ldots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_n)_j = \alpha \cdot \prod_{j \in \{1, \ldots, n\}} (x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)_{\text{ij}} + \beta \cdot \prod_{j \in \{1, \ldots, n\}} (x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)_j \text{ for all } i \in \{1, \ldots, n\}\right\}$ then we have :

1. If $n = 1$ then $\prod_{i \in \{1, \ldots, 1\}} (x_1, \ldots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_n) = \alpha \cdot x + \beta \cdot y = \alpha \cdot \prod_{i \in \{1, \ldots, 1\}} (x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)_i + \beta \cdot \prod_{i \in \{1, \ldots, 1\}} (x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)$ so we have $1 \in S$
2. Let $n \in S$ then for $n + 1$ we have for $i \in \{1, \ldots, n + 1\}$ the following cases
    a. **($\boldsymbol{i \in \{1, \ldots, n\}}$)** So that we have $\prod_{j \in \{1, \ldots, n+1\}} (x_1, \ldots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_{n+1})_j = \left(\prod_{j \in \{1, \ldots, n\}} (x_1, \ldots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_{n+1})_j\right) \cdot x_{n+1} = \left(\prod_{j \in \{1, \ldots, n\}} (x_1, \ldots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_n)_j\right) \cdot x_{n+1} \underset{n \in \mathcal{S}}{=} \left(\alpha \cdot \prod_{j \in \{1, \ldots, n\}} (x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)_j + \beta \cdot \prod_{j \in \{1, \ldots, n\}} (x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)_j\right) \cdot x_{n+1} = \alpha \cdot \left(\prod_{j \in \{1, \ldots, n\}} (x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)_j\right) \cdot x_{n+1} + \beta \cdot \left(\prod_{j \in \{1, \ldots, n\}} (x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)_j\right) \cdot x_{n+1} = \alpha \cdot \left(\prod_{j \in \{1, \ldots, n\}} (x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_{n+1})_j\right) \cdot x_{n+1} + \beta \cdot \left(\prod_{j \in \{1, \ldots, n\}} (x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_{n+1})_j\right) \cdot x_{n+1} = \alpha \cdot \prod_{j \in \{1, \ldots, n+1\}} (x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_{n+1})_j + \beta \cdot \prod_{j \in \{1, \ldots, n+1\}} (x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_{n+1})_j$ so we have $n + 1 \in \mathcal{S}$
    b. **($\boldsymbol{i = n + 1}$)** Then we have that $\prod_{j \in \{1, \ldots, n+1\}} (x_1, \ldots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_{n+1})_j = \left(\prod_{j \in \{1, \ldots, n+1\}} (x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_{n+1})_j\right) \cdot (\alpha \cdot x + \beta \cdot y) = \alpha \cdot \left(\prod_{j \in \{1, \ldots, n+1\}} (x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_{n+1})_j\right) \cdot x + \beta \cdot \left(\prod_{j \in \{1, \ldots, n+1\}} (x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_{n+1})_j\right) \cdot y = \alpha \cdot \prod_{j \in \{1, \ldots, n+1\}} (x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_{n+1})_j + \beta \cdot \prod_{j \in \{1, \ldots, n+1\}} (x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_{n+1})_j$ so that $n + 1 \in \mathcal{S}$

Using mathematical induction we have that $\mathcal{S} = \mathbb{N}$ proving our theorem. □

Example 10.231. Let $n \in \mathbb{N}$ and $\{X_i\}_{i \in \{1, \ldots, n\}}$ a family of vector spaces over a field F and $\{L_i\}_{i \in \{1, \ldots, n\}}$ a family in $\bigcup_{i \in \{1, \ldots, n\}} \operatorname{Hom}(X_i, F)$ such that $\forall i \in \{1, \ldots, n\}$ we have $L_i \in \operatorname{Hom}(X_i, F)$ then if we define $\prod_{i \in \{1, \ldots, n\}} L_i: \prod_{i \in I} X_i \rightarrow F$ by $(x_1, ., x_n) \rightarrow \left(\prod_{i \in \{1, \ldots, n\}} L_i\right)(x_1, \ldots, x_n) = \prod_{i \in \{1, \ldots, n\}} L_i(x_i)$ is multilinear

Proof. Let $(x_1, \ldots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_n) \in \prod_{i \in \{1, \ldots, n\}} X_i$ then $\left(\prod_{j \in \{1, \ldots, n\}} L_j\right)(x_1, \ldots x_{i-1}, \alpha \cdot x + \beta \cdot y, x_i, \ldots, x_n) = \prod_{j \in \{1, \ldots, n\}} (L(x_1), \ldots, L(x_{i-1}), L(\alpha \cdot x + \beta \cdot y), L(x_{i+1}), \ldots, L(x_n))_j = \prod_{j \in \{1, \ldots, n\}} (L(x_1), \ldots, L(x_{i-1}), \alpha \cdot L(x) + \beta \cdot L(y), L(x_{i+1}), \ldots, L(x_n))_j \underset{\text{previous theorem}}{=} \alpha \cdot \prod_{j \in \{1, \ldots, n\}} (L(x_1), \ldots, L(x_{i-1}), L(x), L(x_{i+1}), \ldots, L(x_n))_j + \beta \cdot \prod_{j \in \{1, \ldots, n\}} (L(x_1), \ldots, L(x_{i-1}), L(y), L(x_{i+1}), \ldots, L(x_n))_j = \alpha \cdot \left(\left(\prod_{i \in \{1, \ldots, n\}} L_i\right)(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)\right) + \beta \cdot \left(\left(\prod_{i \in \{1, \ldots, n\}} L_i\right)(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)\right)$ □

Definition 10.232. *Let $n \in \mathbb{N}$ and $\{X_i\}_{i \in \{1, \ldots, n\}}$ a family of vector spaces over a field F, Y a vector space over a field F then* $\mathrm{Hom}(X_1, \ldots, X_n; Y) = \left\{L \in Y^{\prod_{i \in \{1, \ldots, n\}} X_i} | L \text{ is multilinear}\right\}$ *[So* $\mathrm{Hom}(X_1, \ldots, X_n; Y)$ *is the set of multilinear mappings* $L: \prod_{i \in \{1, \ldots, n\}} X_i \to Y$*].*

Theorem 10.233. *Let $n \in \mathbb{N}$ and $\{X_i\}_{i \in \{1, \ldots, n\}}$ a family of vector spaces over a field F, Y a vector space over a field F then* $\mathrm{Hom}(X_1, \ldots, X_n; Y)$ *is a sub space of* $Y^{\prod_{i \in \{1, \ldots, n\}} X_i}$*. Further as* $0 \in \mathrm{Hom}(X_1, \ldots, X_n, Y)$ *it follows that* $\mathrm{Hom}(X_1, \ldots, X_n; Y)$ *is a vector space.*

Proof. If $\alpha, \beta \in F$ and $L_1, L_2 \in \mathrm{Hom}(X_1, \ldots, X_n; Y)$ then we have to prove that $\alpha \cdot L_1 + \beta \cdot L_2 \in \mathrm{Hom}(X_1, \ldots, X_n, Y)$. So if $\alpha', \beta' \in F$, $i \in \{1, \ldots, n\}$, $y, z \in X_i$ and $(x_1, \ldots, x_{i-1}, \alpha' \cdot y + \beta' \cdot z, x_{i+1}, \ldots, x_n)$ then $(\alpha \cdot L_1 + \beta \cdot L_2)(x_1, \ldots, x_{i-1}, \alpha' \cdot y + \beta' \cdot z, x_{i+1}, \ldots, x_n) = \alpha \cdot L_1(x_1, \ldots, x_{i-1}, \alpha' \cdot y + \beta' \cdot z, x_{i+1}, \ldots, x_n) + \beta \cdot L_2(x_1, \ldots, x_{i-1}, \alpha' \cdot y + \beta' \cdot z, x_{i+1}, \ldots, x_n) \underset{L_1, L_2 \text{ is multilinear}}{=} \alpha \cdot (\alpha' \cdot L_1(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) + \beta' \cdot L_1(x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_n)) + \beta \cdot (\alpha' \cdot L_2(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) + \beta' \cdot L_2(x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_n)) = \alpha' \cdot (\alpha \cdot L_1(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) + \beta \cdot L_2(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)) + \beta' \cdot (\alpha \cdot L_1(x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_n) + \beta \cdot L_2(x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_n)) = \alpha' \cdot (\alpha \cdot L_1 + \beta \cdot L_2)(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) + \beta' \cdot (\alpha \cdot L_1 + \beta \cdot L_2)(x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_n)$ proving that $\alpha \cdot L_1 + \beta \cdot L_2$ is multilinear and thus element of $\mathrm{Hom}(X_1, \ldots, X_n, Y)$. □

Theorem 10.234. *Let $n \in \mathbb{N}$ and $\{X_i\}_{i \in \{1, \ldots, n\}}$ a family of vector spaces over a field F, Y, Z vector spaces over a field F then if* $K \in \mathrm{Hom}(X_1, \ldots, X_n; Y)$ *is a multilinear mapping and* $L \in \mathrm{Hom}(Y, Z)$ *a linear mapping then* $L \circ K$ *is a multilinear mapping (so* $L \circ K \in \mathrm{Hom}(X_1, \ldots, X_n; Z)$

Proof. Let $\alpha, \beta \in F$, $i \in \{1, \ldots, n\}$ $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots x_n) \in \prod_{j \in \{1, \ldots, i-1, i+1, .., n\}} X_j$, $x, y \in X_i$ then we have $(L \circ K)(x_1, \ldots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_n) = L(K(x_1, \ldots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_n)) \underset{K \text{ is multilinear}}{=} L(\alpha \cdot K(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) + \beta \cdot K(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)) \underset{L \text{ is linear}}{=} \alpha \cdot L(K(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)) + \beta \cdot L(K(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)) = \alpha \cdot (L \circ K)(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) = \beta \cdot (L \circ K)(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)$. □

Definition 10.235. *Let $n \in \mathbb{N}$ and $\{X_i\}_{i \in \{1, \ldots, n+1\}}$ a family of vector spaces over a field F, then if* $x \in \prod_{i \in \{1, \ldots, n\}} X_i$ *and* $a \in X_{n+1}$ *we define* $x_{\ldots a} \in \prod_{i \in \{1, \ldots, n+1\}} X_i$ *by* $(x_{\ldots a})_k = \begin{cases} x_k \text{ if } k \in \{1, \ldots, n\} \\ a \text{ if } k = n+1 \end{cases}$ *[in other words we have that* $x_{\ldots a} = (x_1, \ldots, x_n, a)$*]*

Theorem 10.236. *Let $n\in\mathbb{N}$ and $\{X_i\}_{i\in\{1,\ldots,n+1\}}$ a family of vector spaces over a field F, then if $x\in\prod_{i\in\{1,\ldots,n\}}X_i$ and $a\in X_{n+1}$ then if $i\in\{1,\ldots,n\}$ and $b\in X_i$ then we have that $(x_{\ldots a})_{j\rightarrow b}=(x_{j\rightarrow b})_{\ldots a}$*

Proof. Let $k\in\{1,\ldots n+1\}$ then $[(x_{\ldots a})_{j\rightarrow b}]_k=\left\{\begin{smallmatrix}(x_{\ldots a})_k \text{ if } k\in\{1,\ldots,n+1\}\setminus\{j\}\\ b \text{ if } k=j\end{smallmatrix}\right.=\left\{\begin{smallmatrix}x_k \text{ if } k\in\{1,\ldots n\}\setminus\{j\}\\ a \text{ if } k=n+1\\ b \text{ if } k=j\end{smallmatrix}\right.=\left\{\begin{smallmatrix}(x_{j\rightarrow b})_k \text{ if } k\in\{1,\ldots n\}\\ a \text{ if } k=n+1\end{smallmatrix}\right.=[(x_{j\rightarrow b})_{\ldots a}]_k$ so that $(x_{\ldots a})_{j\rightarrow b}=(x_{j\rightarrow b})_{\ldots a}$. □

Definition 10.237. *Let $n\in\mathbb{N}$ and $\{X_i\}_{i\in\{1,\ldots,n+1\}}$ a family of vector spaces over a field F, Y a vector space over F then if $L\colon\prod_{i\in\{1,\ldots,n+1\}}X_i\rightarrow Y$ is a mapping and $a\in X_{n+1}$ then we define $L_{(\ldots a)}\colon\prod_{i\in\{1,\ldots,n\}}X_i\rightarrow Y$ by $x\rightarrow L_{(\ldots a)}(x_{\ldots a})$ [in other words $L_{(\ldots a)}(x_1,\ldots,x_n)=L(x_1,\ldots,x_n,a)$]*

Theorem 10.238. *Let $n\in\mathbb{N}$ and $\{X_i\}_{i\in\{1,\ldots,n+1\}}$ a family of vector spaces over a field F, Y a vector space over F then if $L\colon\prod_{i\in\{1,\ldots,n+1\}}X_i\rightarrow Y$ is a multilinear mapping and $a\in X_{n+1}$ then $L_{(\ldots a)}\colon\prod_{i\in\{1,\ldots,n\}}X_i\rightarrow Y$ is multilinear.*

Proof. Take $i\in\{1,\ldots,n\}$, $\alpha,\beta\in F$, $y,z\in X_i$ and $x\in\prod_{i\in\{1,\ldots,n\}}X_i$ such that $x_i=\alpha\cdot y+\beta\cdot z$ then for $x_{\ldots a}\in\prod_{i\in\{1,\ldots,n+1\}}$ we have that $(x_{\ldots a})_i\underset{i\neq n+1}{=}x_i=\alpha\cdot y+\beta\cdot z$ so that we have that $L_{(\ldots a)}(x)=L(x_{\ldots a})\underset{\text{L is multilinear}}{=}\alpha\cdot L((x_{\ldots a})_{i\rightarrow y})+\beta\cdot L((x_{\ldots a})_{i\rightarrow z})\underset{10.236}{=}\alpha\cdot L((x_{i\rightarrow y})_{\ldots a})+\beta\cdot L((x_{i\rightarrow z})_{\ldots a})=\alpha\cdot L_{(\ldots a)}(x_{i\rightarrow y})+\beta\cdot L_{(\ldots a)}(x_{i\rightarrow z})$ □

If we use the notation that is introduced in 10.225 then the proof becomes almost trivial as seen in the following version of the above proof. From now on we try to use the easy notation on our proofs.

Proof. Take $i\in\{1,\ldots,n\}$, $\alpha,\beta\in F$, $y,z\in X_i$ and $x=(x_1,\ldots,x_{i-1},\alpha\cdot y+\beta\cdot z,x_{i+1},\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}X_i$ then $L_{(\ldots a)}(x_1,\ldots,x_{i-1},\alpha\cdot y+\beta\cdot z,x_{i+1},\ldots,x_n)=L(x_1,\ldots,x_{i-1},\alpha\cdot y+\beta\cdot x,x_{i+1},\ldots,x_n,a)\underset{L\,\text{is}\,\text{multilinear}}{=}\alpha\cdot L(x_1,\ldots,x_{i-1},y,x_{i+1},\ldots,x_n,a)+\beta\cdot L(x_1,\ldots,x_{i-1},z,x_{i+1},\ldots,x_n,a)=\alpha\cdot L_{(\ldots a)}(x_1,\ldots,x_{i-1},y,x_{i+1},\ldots,x_n)+\beta\cdot L(x_1,\ldots,x_{i-1},z,x_{i+1},\ldots,x_n)$ □

The above theorem can be used in the induction step of proofs like the following theorem will show.

Theorem 10.239. *Let $n\in\mathbb{N}$, $\{X_i\}_{i\in\{1,\ldots,n\}}$ a family of vector spaces over a field F, Y a vector space over Y and $\alpha=(\alpha_1,\ldots,\alpha_n)\in F^n$, $x=(x_1,\ldots,x_n)\in\prod_{i\in[1,\ldots,n]}X_i$, $L\colon\prod_{i\in\{1,\ldots,n\}}X_i\rightarrow Y$ a multilinear function then if we define $\alpha\cdot x\in\prod_{i\in\{1,\ldots,n\}}X_i$ by $\forall i\in\{1,\ldots,n\}$ $(\alpha\cdot x)_i=\alpha_i\cdot x_i$ or $\alpha\cdot z=(\alpha_1\cdot x_1,\ldots,\alpha_n\cdot x_n)$ we have that $L(\alpha\cdot x)=\left(\prod_{i\in\{1,\ldots,n\}}\alpha_i\right)\cdot L(x)$*

Proof. We prove this by induction on n so let $S=\left\{n\in\{1,\ldots,n\}|\text{If } L: \prod_{i\in\{1,\ldots,n\}} X_i\rightarrow Y \text{ is multilinear and } \alpha\in F^n, x\in\prod_{i\in\{1,\ldots,n\}} X_i \text{ then } L(\alpha\cdot x)=\left(\prod_{i\in\{1,\ldots,n\}}\alpha_i\right)\cdot L(x)\right\}$ then we have:

1. If $n=1$ then $\alpha=(\alpha_1), x=(x_1)$ and $L(\alpha\cdot x)=L(\alpha_1\cdot x_1)=\alpha_1\cdot L(x_1)=\left(\prod_{i\in\{1\}}\alpha_i\right)\cdot L(x_1)=\left(\prod_{i\in\{1\}}\alpha_i\right)\cdot L(x)$ so $1\in S$

2. If $n\in S$ take then $\alpha=(\alpha_1,\ldots,\alpha_{n+1})\in F^{n+1}$, $x\in\prod_{i\in\{1,\ldots,n+1\}} X_i$, $L: \prod_{i\in\{1,\ldots,n+1\}} X_i\rightarrow Y$ a multilinear mapping then as trivially $\alpha\cdot x=(\alpha\cdot x_{|\{1,\ldots,n\}})_{(\ldots\alpha_{n+1}\cdot x_{n+1})}$ we have $L(\alpha\cdot x)=L((\alpha\cdot x_{|\{1,\ldots,n\}})_{\ldots\alpha_{n+1}\cdot x_{n+1}})=L_{(\ldots\alpha_{n+1}\cdot x_{n+1})}(\alpha\cdot x_{|\{1,\ldots,n\}|})\underset{n\in S \text{ and } 10.238}{=}\left(\prod_{i\in\{1,\ldots,n\}}\alpha_i\right)\cdot L_{(\ldots\alpha_{+1}\cdot x_{n+1})}(x_{|n+1})=\left(\prod_{i\in\{1,\ldots,n\}}\alpha_i\right)\cdot L(x_1,\ldots,x_n,\alpha_{n+1}\cdot x_{n+1})\underset{\text{L is multilinear}}{=}\left(\prod_{i\in\{1,\ldots,n\}}\alpha_i\right)\cdot(\alpha_{n+1}\cdot L(x_1,\ldots,x_{n+1}))=\left(\left(\prod_{i\in\{1,\ldots,n\}}\alpha_i\right)\cdot\alpha_{n+1}\right)\cdot L(x)=\left(\prod_{i\in\{1,\ldots,n+1\}}\alpha_i\right)\cdot L(x)$ proving that $n+1\in S$

Using mathematical induction we have then that $S=\{1,\ldots\}=\mathbb{N}$ proving the theorem. □

Theorem 10.240. *Let $n\in\mathbb{N}$, $\{X_i\}_{i\in\{1,\ldots,n\}}$ a family of vector spaces over a field F, Y a vector-space over F and $L: \prod_{i\in\{1,\ldots,n\}} X_i\rightarrow Y$ is a multi-linear function then if $x\in\prod_{i\in\{1,\ldots,n\}} X_i$ is such that $\exists i\in\{1,\ldots,n\}$ such that $x_i=0$ then $L(x)=0$*

Proof. $L(x_1,\ldots,x_{i-1},0,x_i,\ldots,x_n)=L(x_1,\ldots,x_{i-1},0\cdot 0,x_{i+1},\ldots,x_n)=0\cdot L(x_1,\ldots,x_{i-1},0,x_{i+1},\ldots,x_n)=0$ □

Theorem 10.241. *Let $n\in\mathbb{N}$, $\{X_i\}_{i\in\{1,\ldots,n\}}$ a family of vector spaces over a field F, $i\in\{1,\ldots,n\}$, $a\in X_i$ and $\sigma\in P_n$ then $(x_1,\ldots,x_{i-1},a,x_{i+1},\ldots,x_n)\circ\sigma=((x\circ\sigma)_1,\ldots,(x\circ\sigma)_{\sigma^{-1}(i)-1},a,(x\circ\sigma)_{\sigma^{-1}(i)+1},\ldots,(x\circ\sigma)_{n+1})$ (in other words if $x_i=a$ then $(x\circ\sigma)_{\sigma^{-1}(i)}=a$)*

Proof. If $x=(x_1,\ldots,x_{i-1},a,x_{i+1},\ldots,x_n)$ then this is by definition equivalent to $x_i=a$, then we have $(x\circ\sigma)_{\sigma^{-1}(i)}=x(\sigma(\sigma^{-1}(i)))=x(i)=x_i=a$. □

Theorem 10.242. *Let Y be a vector space over a field F, $n\in\mathbb{N}$, $\{X_i\}_{i\in\{1,\ldots,n\}}$ a family of vector spaces over F and then*

1. *If $m\in\mathbb{N}$ and $\{L_i\}_{i\in\{0,\ldots,m\}}$ is a family in $\mathrm{Hom}(X_1,\ldots,X_n;Y)$ then if we define $\sum_{i=0}^m L_i: \prod_{i\in\{1,\ldots,n\}} X_i\rightarrow Y$ by $x\rightarrow\sum_{i=0}^m L_i(x)$ we have that $\sum_{i=0}^m L_i$ is a multilinear function.*

2. *If $\{L_i\}_{i\in I}$ is a finite family (meaning I is finite) in $\mathrm{Hom}(X_1,\ldots,X_n;Y)$ then if we define $\sum_{i\in I} L_i|\prod_{i\in\{1,\ldots,n\}} X_i\rightarrow Y$ by $x\rightarrow\sum_{i\in I} L_i(x)$ (which is defined as I is finite) we have that $\sum_{i\in I} L_i$ is multilinear.*

Proof.

1. Let $\alpha,\beta\in F$, $i\in\{1,\ldots,n\}$, $y,z\in X_i$ and $x=(x_1,\ldots,x_{i-1},\alpha\cdot y+\beta\cdot z,x_{i+1},\ldots,x_n)\in\prod_{j\in\{1,\ldots,n\}}X_i$ then $(\sum_{j=0}^m L_j)(x)=\sum_{j=0}^m L_j(x) \underset{L_i\,\text{are multilinear}}{=} \sum_{j=0}^m(\alpha\cdot L_j(x_{i\rightarrow y})+\beta\cdot L_j(x_{i\rightarrow z}))=\sum_{j=0}^m \alpha\cdot L_j(x_{i\rightarrow y})+\sum_{j=0}^m\beta\cdot L_j(x_{i\rightarrow z})=\alpha\cdot\sum_{j=0}^m L_j(x_{i\rightarrow y})+\beta\cdot\sum_{j=0}^m L_j(x_{i\rightarrow z})=\alpha\cdot(\sum_{j=0}^m L_j)(x_{i\rightarrow y})+\beta\cdot(\sum_{j=0}^m L_j)(x_{i\rightarrow z})$

2. Let $\alpha,\beta\in F$, $i\in\{1,\ldots,n\}$, $y,z\in X_i$ and $x=(x_1,\ldots,x_{i-1},\alpha\cdot y+\beta\cdot z,x_{i+1},\ldots,x_n)\in\prod_{j\in\{1,\ldots,n\}}X_i$ then $(\sum_{j\in I} L_j)(x)=\sum_{j\in I} L_j(x) \underset{L_i\,\text{are multilinear}}{=} \sum_{j\in I}(\alpha\cdot L_j(x_{i\rightarrow y})+\beta\cdot L_j(x_{i\rightarrow z}))=\sum_{j\in I}\alpha\cdot L_j(x_{i\rightarrow y})+\sum_{j\in I}\beta\cdot L_j(x_{i\rightarrow z})=\alpha\cdot\sum_{j\in I}L_j(x_{i\rightarrow y})+\beta\cdot\sum_{j\in I}L_j(x_{i\rightarrow z})=\alpha\cdot(\sum_{j\in I}L_j)(x_{i\rightarrow y})+\beta\cdot(\sum_{j\in I}L_j)(x_{i\rightarrow z})$ □

Theorem 10.243. *Let $n\in\mathbb{N}$, $\{X_i\}_{i\in\{1,\ldots,n\}}$ a family of vector spaces over a field F, Y a vector space over a field F then if $L\in\operatorname{Hom}(X_1,\ldots,X_n;Y)$ and $i\in\{1,\ldots,n\}$:*

1. *If $m\in\mathbb{N}$ and $\{\alpha_j\}_{j\in\{1,\ldots,m\}}$ is a family in F and $\{y_j\}_{j\in\{1,\ldots,m\}}$ a family in X_i then if $x=(x_1,\ldots,x_{i-1},\sum_{j=1}^m\alpha_j\cdot x_j,x_{i+1},\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}X_i$ we have $L(x_1,\ldots,x_{i-1},\sum_{j=1}^m\alpha_j\cdot x_j,x_{i+1},\ldots,x_n)=\sum_{j=1}^m\alpha_jL(x_1,\ldots,x_{i-1},y_j,x_{i+1},\ldots,x_n)$*

2. *If $\{\alpha_j\}_{j\in I}$ is a family in F with finite support and $\{y_j\}_{j\in I}$ a family in X_i then if $x=(x_1,\ldots,x_{i-1},\sum_{j\in I}\alpha_j\cdot x_j,x_{i+1},\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}X_i$ we have $L(x_1,\ldots,x_{i-1},\sum_{j\in I}\alpha_j\cdot x_j,x_{i+1},\ldots,x_n)=\sum_{j\in I}\alpha_jL(x_1,\ldots,x_{i-1},y_j,x_{i+1},\ldots,x_n)$*

Proof.

1. We prove this by induction on m so let $S_i=\{m\in\{1,\ldots\}|\text{if }\{\alpha_j\}_{j\in\{1,\ldots,m\}},\{y_j\}_{j\in\{1,\ldots,m\}}\text{ are families in }F,X_i\text{ then if }(x_1,\ldots,x_{i-1},\sum_{j=1}^m\alpha_j\cdot y_j,x_{i+1},\ldots,x_n)\text{ we have }L(x_1,\ldots,x_{i-1},\sum_{j=1}^m\alpha_j\cdot y_j,x_{i+1},\ldots,x_n)=\sum_{j=1}^m\alpha_j\cdot L(x_1,\ldots,x_{i-1},y_j,x_{i+1},\ldots,x_n)\}$ then we have:

   a. If $m=1$ then if $\{\alpha_j\}_{j\in\{1\}},\{y_j\}_{j\in\{1\}}$ are families in F,X_i then $L(x_1,\ldots,x_{i-1},\sum_{j=1}^1\alpha_j\cdot x_j,x_{i+1},\ldots,x_n)=L(x_1,\ldots,x_{i-1},\alpha_1\cdot x_1,x_{i+1},\ldots,x_n)=\alpha_1\cdot L(x_1,\ldots,x_{i-1},y_1,x_{i+1},\ldots,x_n)=\sum_{j=1}^1\alpha_j\cdot L(x_1,\ldots,x_{i-1},y_j,x_{i+1},\ldots,x_n)$ proving that $1\in S_i$.

   b. If $m\in S_i$ then if $\{\alpha_j\}_{j\in\{1,\ldots,m+1\}}$, $\{y_j\}_{j\in\{1,\ldots,m+1\}}$ are families inf F,X_i then $L(x_1,\ldots,x_{i-1},\sum_{j=1}^{m+1}\alpha_j\cdot y_j,x_{i+1},\ldots,x_n)=L(x_1,\ldots,x_{i-1},\sum_{j=1}^m\alpha_j\cdot y_j+\alpha_{m+1}\cdot y_{m+1},x_{i+1},\ldots,x_n)=L(x_1,\ldots,x_{i-1},\sum_{j=1}^m\alpha_j\cdot y_j,x_{i+1},\ldots,x_n)+\alpha_{m+1}\cdot L(x_1,\ldots,x_{i-1},y_{m+1},x_{i+1},\ldots,x_n)\underset{m\in S_i}{=}\sum_{j=1}^m\alpha_j\cdot L(x_1,\ldots,x_{i-1},y_j,x_{i+1},\ldots,x_n)+\alpha_{m+1}\cdot L(x_1,\ldots,x_{i-1},y_{m+1},x_{i+1},\ldots,x_n)=\sum_{j=1}^{m+1}\alpha_j\cdot L(x_1,\ldots,x_{i-1},y_j,x_{i+1},\ldots,x_n)$ proving that $m+1\in S_i$.

   Using mathematical induction we have that $S_i=\{1,\ldots,\}=\mathbb{N}$

2. Let $\{\alpha_j\}_{j\in I}$ be a family in F with finite support, $\{y_j\}_{j\in I}$ a family in X_i. We have $\operatorname{support}(\{\alpha_j\cdot x_j\}_{j\in I})\subseteq\operatorname{support}(\{\alpha_j\}_{j\in I})$ (see 10.126). Now if $k\in\operatorname{support}(\{\alpha_j\cdot L(x_1,\ldots,x_{i-1},y_j,x_{i+1},\ldots,x_n)\}_{j\in I})\Rightarrow\alpha_k\cdot L(x_1,\ldots,x_{i-1}, y_k, x_{i+1},\ldots,x_n)\neq 0\Rightarrow\alpha_k\neq 0\Rightarrow k\in\operatorname{support}(\{\alpha_j\}_{j\in I})$, so $\operatorname{support}(\{\alpha_j\cdot x_j\}_{j\in I})$,$\operatorname{support}(\{\alpha_j\cdot L(x_1,\ldots,x_{i-1},y_j,x_{i+1},\ldots,x_n)\}_{j\in I})\subseteq\operatorname{support}(\{\alpha_j\}_{j\in I})$ proving that $\operatorname{support}(\{\alpha_j\cdot x_j\}_{j\in I})$,$\operatorname{support}(\{\alpha_j\cdot L(x_1,\ldots,x_{i-1},y_j,x_{i+1},\ldots, x_n)\}_{j\in I})$ is finite and thus that both sums in (2) are well defined. If now $b:\{0,\ldots,m\}\to\operatorname{support}(\{\alpha_i\}_{i\in I})$ then we have $L(x_1,\ldots,x_{i-1},\sum_{j\in I}\alpha_j\cdot y_j, x_{i+1},\ldots,x_n)=L(x_1,\ldots,x_{i-1},\sum_{j=0}^m\alpha_{b_j}\cdot y_{b_j},x_{i+1},\ldots,x_n)=L(x_1,\ldots,x_{i-1}, \sum_{j=1}^{m+1}\alpha_{b_{j-1}}\cdot y_{b_{j-1}},x_{i+1},\ldots,x_n)\underset{(1)}{=}\sum_{i=1}^{m+1}\alpha_{b_{j-1}}\cdot L(x_1,\ldots,x_{i-1},y_{b_{j-1}},x_{i+1},\ldots, x_n)=\sum_{j=0}^m\alpha_{b_j}\cdot L(x_1,\ldots,x_{i-1},y_{b_j},x_{i+1},\ldots,x_n)=\sum_{j\in I}\alpha_j\cdot L(x_1,\ldots,x_{i-1}, y_j,x_{i+1},\ldots,x_n)$ □

The following theorem proves that given a family of sub spaces of a family of vectorspaces we can always extend multilinear mappings on the sub spaces to multilinear mappings on the vector space.

Theorem 10.244. *Let $n\in\mathbb{N}$ and $\{X_i\}_{i\in\{1,\ldots,n\}}$ a family of vector spaces over a field F, Y a vector space over F and let $\{Y_i\}_{i\in\{1,\ldots,n\}}$ a family of non empty sub spaces [i.e. $\forall i\in\{1,\ldots,n\}$ we have that Y_i is a non empty subspace of X_i]. If $L\in\operatorname{Hom}(Y_1,\ldots,Y_n;Y)$ then we can find a $K\in\operatorname{Hom}(X_1,\ldots,X_n;Y)$ such that $K_{|\prod_{i\in\{1,\ldots,n\}}Y_i}=L$*

Proof. Using 10.194 we have $\forall i\in\{1,\ldots,n\}$ that there exists a subspace Z_i such that $X_i=Y_i\oplus Z_i$ and thus $\forall x\in X_i\ \exists!y\in Y_i\wedge\exists!z\in Z_i$ such that $x=y+z$, this defines two functions $\tau_i: X_i\to Y_i$ and $\rho_i: X_i\to Z_i$ such that $x=\tau_i(x)+\rho_i(x)$. Note that if $\alpha,\beta\in F$ and $x,y\in X_i$ then we have $x=\tau_i(x)+\rho_i(x)$, $y=\tau_i(y)+\rho_i(y)$ then $\alpha\cdot x+\beta\cdot y=\alpha\cdot(\tau_i(x)+\rho_i(x))+\beta\cdot(\tau_i(y)+\rho_i(y))=(\alpha\cdot\tau_i(x)+\beta\cdot\tau_i(y))+(\alpha\cdot\rho_i(x)+\beta\cdot\rho_i(y))\underset{\text{extension is unique in direct sum}}{\Rightarrow}\tau_i(\alpha\cdot x+\beta\cdot y)=\alpha\cdot\tau_i(x)+\beta\cdot\tau_i(y)\wedge \rho_i(\alpha\cdot x+\beta\cdot y)=\alpha\cdot\rho_i(x)+\beta\cdot\rho_i(y)$ proving that τ_i, ρ_i are linear. Given now L define $K:\prod_{i\in\{1,\ldots,n\}}X_i\to Y$ by $(x_1,\ldots,x_n)\to K(x_1,\ldots,x_n)=L(\tau_1(x_1),\ldots,\tau_n(x_n))$. We have now $K(x_1,\ldots,x_{i-1},\alpha\cdot x+\beta\cdot y,x_{i+1},\ldots,x_n)=L(\tau_1(x_1),\ldots,\tau_{i-1}(x_{i-1}),\tau_i(\alpha\cdot x+\beta\cdot y),\tau_{i+1}(x_{i+1}),\ldots,\tau_n(x_n))\underset{\tau_i\text{ is linear}}{=}L(\tau_1(x_1),\ldots,\tau_{i-1}(x_{i-1}),\alpha\cdot\tau_i(x)+\beta\cdot\tau_i(y), \tau_{i+1}(x_{i+1}),\ldots,\tau_n(x_n))\underset{L\text{ is multilinear}}{=}\alpha\cdot L(\tau_1(x_1),\ldots,\tau_{i-1}(x_{i-1}),\tau_i(x),\tau_{i+1}(x_{i+1}),\ldots, \tau_n(x_n))+\beta\cdot L(\tau_1(x_1),\ldots,\tau_{i-1}(x_{i-1}),\tau_i(y),\tau_{i+1}(x_{i+1}),\ldots,\tau_n(x_n))=\alpha\cdot K(x_1,\ldots,x_{i-1}, x,x_{i+1},\ldots,x_n)+\beta\cdot K(x_1,\ldots,x_{i-1},y,x_{i+1},\ldots,x_n)$ proving that K is multilinear or $K\in\operatorname{Hom}(X_1,\ldots,X_n;Y)$.

If now $(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}Y_i\subseteq\prod_{i\in\{1,\ldots,n\}}X_i$ then $\forall i\in\{1,\ldots,n\}$ we have $x_i=x_i+0\underset{x_i\in Y_i,0\in Z_i}{\Rightarrow}\tau_i(x_i)=x_i$ and thus $K(x_1,\ldots,x_n)=L(\tau_1(x_1),\ldots,\tau_n(x_n))=L(x_1,\ldots, x_n)$ proving that $K_{|\prod_{i\in\{1,\ldots,n\}}Y_i}=L$ □

10.8 Determinant Functions

First we have to introduce the concept of -1 which we need in addition to 1 for the definition of skew-symmetric functions.

Definition 10.245. *If F is a field and 1 is the multiplicative unit then -1 is the (sum) inverse of 1 so -1 is such that $(-1)+1=0=1+(-1)$*

We have then the following theorem for a vector space

Theorem 10.246. *If X is a vector field over F then we have that $\forall x\in X$ that $-x=(-1)\cdot x$*

Proof. We have $(-1)\cdot x+x=(-1)\cdot x+1\cdot x=((-1)+1)\cdot x=0\cdot x=0=0\cdot x=(1+(-1))\cdot x=1\cdot x+(-1)\cdot x=x+(-1)\cdot x$ □

Next we have a generalization of power's in $\mathbb{Z}$

Definition 10.247. *If F is a field then as $\langle F,\cdot\rangle$ is a abelian semi-group we have by 4.22 that given a $a\in F$ and $n\in\mathbb{N}_0$ that there exists a a^n such that*

$$\begin{aligned} a^0 &= 1 \\ a^{n+1} &= a^n\cdot a \underset{\text{abelian}}{=} a\cdot a^n \end{aligned}$$

Theorem 10.248. *If $n,n'\in\mathbb{N}_0$ and $a\in F$ then $a^{n'+n}=a^{n'}\cdot a^n$*

Proof. We prove this by induction on n. So let $X=\{n\in\mathbb{N}_0|a^{n'+n}=a^{n'}\cdot a^n\}$ then we have

1. If $n=0$ then $a^{n'+n}=a^{n'+0}=a^{n'}=a^{n'}\cdot 1=a^{n'}\cdot a^0\Rightarrow 0\in X$
2. If $n\in X$ then $a^{n'+(n+1)}=a^{(n'+n)+1}=a^{(n'+n)}\cdot a \underset{n\in X}{=} (a^{n'}\cdot a^n)\cdot a=a^{n'}\cdot(a^n\cdot a)=a^{n'}\cdot a^{n+1}$ and thus $n+1\in X$

Using mathematical induction (see 4.10) we have $X=\mathbb{N}_0$ proving the theorem □

Theorem 10.249. *In F we have*

$$\begin{aligned} 0^n &= 0\,(\text{if } n\neq 0) \\ 1^n &= 1 \\ (-1)^n &= -1 \text{ or } 1 \\ (-1)^{2\cdot n} &= 1 \\ (-1)^{2\cdot n+1} &= -1 \end{aligned}$$

Proof.

1. If $n\neq 0\Rightarrow\exists m\in\mathbb{N}_0\vdash n=m+1$ then $0^n=0^{(m+1)}=0^m\cdot 0=0$
2. $1^n=1$ is proved by induction on n, let $X=\{n\in\mathbb{N}_0|1^n=1\}$ then
   a. $1^0=1\Rightarrow 0\in X$
   b. If $n\in X\Rightarrow 1^{n+1}=1^n\cdot 1 \underset{n\in X}{=} 1\cdot 1=1\Rightarrow n+1\in X$

   so $X=\mathbb{N}_0$
3. $(-1)^n=\pm 1$ is proved by induction on n, let $X=\{n\in\mathbb{N}_0|(-1)^n=-1\text{ or }1\}$ then
   a. $(-1)^0=1\Rightarrow 0\in X$

b. If $n\in X$ then $(-1)^{n+1}=(-1)^n\cdot(-1)\underset{n\in X}{=}(-1)\cdot(-1)\vee 1\cdot(-1)=1\vee -1\Rightarrow n+1\in X$

so $X=\mathbb{N}_0$

4. $(-1)^{2\cdot n}=(-1)^{(1+1)\cdot n}=(-1)^{n+n}=(-1)^n\cdot(-1)^n\underset{(3)}{=}(-1)\cdot(-1)\,\text{or}\,1\cdot 1=1$
5. $(-1)^{2\cdot n+1}=(-1)^{2\cdot n}\cdot(-1)\underset{(4)}{=}1\cdot(-1)=-1$ □

We can now make identification between $(1,-1)\in\mathbb{Z}$ and $(1,-1)\in F$ such that $(-1)^n$ is the same in $\mathbb{Z}$ and F as is proved in the next theorem.

Theorem 10.250. *If $(1,-1)\in\mathbb{Z}$ and $(1',-1')\in F$ [we use $'$ to show that we are talking about $(1,-1)\in F$] then $\forall n\in\mathbb{N}_0$ we have that $(-1)^n=(-1')^n$ if we equate 1 with $1'$ and -1 with $-1'$*

Proof. This is easily proved by induction on $n\in\mathbb{N}_0$ so let $S=\{n\in\mathbb{N}_0|(-1)^n=(-1')^n\}$ then we have:

1. If $n=0$ then $(-1)^0=1=1'=(-1')^0$ and thus $n\in S$
2. If $n\in S$ then $(-1)^n=(-1')^n$ and we have to consider the following cases

   a. $\boldsymbol{((-1)^n=1=1'=(-1')^n)}$ so $(-1)^{n+1}=(-1)\cdot(-1)^n=(-1)\cdot 1=-1=-1'=(-1')\cdot 1'=(-1')\cdot(-1')^n=(-1')^n$ and thus $n+1\in S$

   b. $\boldsymbol{((-1)^n=-1=-1'=(-1')^n)}$ $(-1)^{n+1}=(-1)\cdot(-1)^n=(-1)\cdot(-1)=1=1'=(-1')\cdot(-1')=(-1')\cdot(-1')^n=(-1')^n$ so $n+1\in S$

   So in all cases we have $n+1\in S$

By mathematical induction we have then $S=\mathbb{N}_0$ proving our theorem. □

For the following definition remember by 10.98 that if $n\in\mathbb{N}$, X is a set and $\{X_i\}_{i\in\{1,\ldots,n\}}$ is a family of sets such that $X_i=X$ then $X^n=\prod_{i\in\{1,\ldots,n\}}X_i$ (which is by 10.99 equal to $X^{\{1,\ldots,n\}}$).

Definition 10.251. *Let $n\in\mathbb{N}$ and X,Y vector spaces over a field F then we say that $L\colon X^n\to Y$ is a n-linear mapping if L is a multilinear mapping. We note $\operatorname{Hom}(X^n;Y)=\{L\in Y^{X^n}|L\,\text{is multilinear}\}$ for the set of n-linear mappings.*

Note 10.252. Using 10.233 we have that $\operatorname{Hom}(X^n,Y)$ is a vector space and a subspace of Y^{X^n}

Definition 10.253. *Let X,Y be vector spaces over the field F, $n\in\mathbb{N}$ then a n-linear mapping $L\in\operatorname{Hom}(X^n;Y)$ is **skew symmetric** if $\forall\sigma\in P_n$ we have $\sigma L=\operatorname{sign}(\sigma)\cdot L$*

Example 10.254. If X,Y are vector spaces over the field F and $n=1$ we have that if $L\in\operatorname{Hom}(X^1;Y)$ then L is skew symmetric

Proof. If $\sigma\in P_1$ then $\sigma=1_{\{1\}}$ with $\operatorname{sign}(\sigma)=1$ and $\forall x\in X^1$ we have $\sigma L(x)=L(x\circ\sigma)=L(x)=\operatorname{sign}(\sigma)\cdot L(x)$ proving that $\sigma L=\operatorname{sign}(\sigma)\cdot L$ □

Lemma 10.255. *Let X, Y be vector spaces over a field F, $n \in \mathbb{N}$ and $L \in \operatorname{Hom}(X^n; Y)$ a n-linear mapping then σL is a multilinear mapping.*

Proof. Let $i \in \{1, \ldots, n\}$, $\alpha, \beta \in F$ $y, z \in X_i$ and $x = (x_1, \ldots, x_{i-1}, \alpha \cdot y + \beta \cdot z, x_{i+1}, \ldots, x_n) \in \prod_{i \in \{1, \ldots, n\}} X_i$ then $\sigma L(x) = L((x_1, \ldots, x_{i-1}, \alpha \cdot y + \beta \cdot z, x_{i+1}, \ldots, x_n) \circ \sigma) \underset{10.241}{=} L(x_{\sigma(1)}, \ldots, x_{\sigma(\sigma^{-1}(i)-1)}, \alpha \cdot y + \beta \cdot z, x_{\sigma(\sigma^{-1}(i)+1)}, \ldots, x_{\sigma(n)}) = \alpha \cdot L(x_{\sigma(1)}, \ldots, x_{\sigma(\sigma^{-1}(i)-1)}, y, x_{\sigma(\sigma^{-1}(i)+1)}, \ldots, x_{\sigma(n)}) + \beta \cdot L(x_{\sigma(1)}, \ldots, x_{\sigma(\sigma^{-1}(i)-1)}, z, x_{\sigma(\sigma^{-1}(i)+1)}, \ldots, x_{\sigma(n)}) = \alpha \cdot L((x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) \circ \sigma) + \beta \cdot L((x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_n) \circ \sigma) = \alpha \cdot \sigma L(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) + \beta \cdot \sigma L(x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_n)$ proving the multilinearity of σL. $\square$

Theorem 10.256. *Let X, Y be vector spaces over a field F then given $n \in \mathbb{N}$ and a n-linear mapping $L \in \operatorname{Hom}(X^n; Y)$ then $\sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \sigma L: X^n \to Y$ defined by $x \to \sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \sigma L$ is skew symmetric (the sum is well defined for we have proved that P_n is a finite set (see 10.70)) and $\sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \sigma L \in \operatorname{Hom}(X^n, Y)$.*

Proof. First as $\operatorname{Hom}(X^n; Y)$ is a vector space and the above lemma we have that $\operatorname{sign}(\sigma) \cdot \sigma L$ is n-linear, then using 10.242 we have that $\sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \sigma L$ is n-linear and thus $\sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \sigma L \in \operatorname{Hom}(X^n; Y)$. Next we have to prove that $\sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \sigma L$ is skew symmetric. Take $\tau \in P_n$ then $\forall x \in X^n$ we have

$$
\begin{aligned}
\tau\left(\sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \sigma L\right)(x) &= \left(\sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \sigma L\right)(x \circ \tau) \\
&= \sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \sigma L(x \circ \tau) \\
&= \sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot L((x \circ \tau) \circ \sigma) \\
&= \sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot L(x \circ (\tau \circ \sigma)) \\
&= \sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot (\tau \circ \sigma) L(x) \\
&\underset{\operatorname{sign}(\tau) \cdot \operatorname{sign}(\tau) = 1}{=} \sum_{\sigma \in P_n} \operatorname{sign}(\tau) \cdot \operatorname{sign}(\tau) \cdot \operatorname{sign}(\sigma) \cdot (\tau \circ \sigma) L(x) \\
&\underset{10.91}{=} \operatorname{sign}(\tau) \cdot \sum_{\sigma \in P_n} \operatorname{sign}(\tau \circ \sigma) \cdot (\tau \circ \sigma) L(x) \\
&\underset{10.97}{=} \operatorname{sign}(\tau) \cdot \sum_{\sigma \in P_n} \operatorname{sign}(T_\tau(\sigma)) \cdot T_\tau(\sigma) L(x) \\
&\underset{10.97 \text{ and } 10.44}{=} \operatorname{sign}(\tau) \cdot \sum_{\sigma \in T_\tau(P_n)} \operatorname{sign}(\sigma) \cdot \sigma L(x) \\
&\underset{10.97}{=} \operatorname{sign}(\tau) \cdot \sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \sigma L(x) \\
&= \operatorname{sign}(\tau) \cdot \left(\sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot L\right)(x)
\end{aligned}
$$

proving that $\tau(\sum_{\sigma\in P_n}\operatorname{sign}(\sigma)\cdot\sigma L)=\operatorname{sign}(\tau)\cdot\sum_{\sigma\in P_n}\operatorname{sign}(\sigma)\cdot\sigma L$ and thus skew symmetry. $\square$

Theorem 10.257. *Let X, Y be vector spaces over a field F that has characteristics of zero (see 4.25) then given $n\in\mathbb{N}$ and a n-linear mapping $L\in\operatorname{Hom}(X^n;Y)$ the the following are equivalent:*

1. *L is skew symmetric*
2. *If $x=(x_1,\ldots,x_n)\in X^n$ is such that $\exists i,j\in\{1,\ldots,n\}$ with $i\neq j$ and $x_i=x_j$ then $L(x)=0$*
3. *If $x=\{x_1,\ldots,x_n\}\in X^n$ is such that $\{x_i\}_{i\in\{1,\ldots,n\}}$ is linear independent then $L(x)=0$*

Proof.

1. **(1 ⇒ 2)** Let $x\in X^n$ with $i,j\in\{1,\ldots,n\}$, $i\neq j$ and $x_i=x_j$ then we have $(x\circ(i\leftrightarrow_n j))_k=\begin{cases}x_j \text{ if } k=i\\ x_i \text{ if } k=j\\ x_k \text{ if } k\in\{1,\ldots,n\}\setminus\{i,j\}\end{cases}\underset{x_i=x_j}{=}\begin{cases}x_i \text{ if } k=i\\ x_j \text{ if } k=j\\ x_k \text{ if } k\in\{1,\ldots,n\}\setminus\{i,j\}\end{cases}=x_k$ proving that $x\circ(i\leftrightarrow_n j)=x$. Now as L is skew symmetric we have that $(i\leftrightarrow_n j)L=\operatorname{sign}(i\leftrightarrow_n j)L$ so that we have (u is unit in F) $-L(x)=\operatorname{sign}(i\leftrightarrow_n j)\cdot L(x)=(i\leftrightarrow_n j)L(x)=L(x\circ(i\leftrightarrow_n j))=L(x)\Rightarrow L(x)+L(x)=0\Rightarrow 0=u\cdot L(x)+u\cdot L(x)=(1+1)\cdot L(x)=(2\cdot u)\cdot L(x)\underset{2\cdot u\neq 0 \text{ and } 10.107}{=}L(x)=0$

2. **(2 ⇒ 1)** Let $x=(x_1,\ldots,x_n)\in X^n$, $i,j\in\{1,\ldots,n\}$ such that $i\neq j$ define then x' by $x'_k=\{x_1,\ldots,x_{i-1},x_i+x_j,x_{i+1},\ldots,x_{j-1},x_i+x_j,x_{j+1},\ldots,x_n\}$ then as $x'_i=x_i+x_j=x'_j$ we have that
$$L(x')=0 \tag{10.61}$$
By using multilinearity we have $L(x')=L(x'_{i\rightarrow x_i})+L(x'_{i\rightarrow x_j})$ giving together with 10.61 that
$$L(x'_{i\rightarrow x_i})+L(x'_{i\rightarrow x_j})=0 \tag{10.62}$$
Using multilinearity of L again we have as $(x'_{i\rightarrow x_i})_j\underset{i\neq j}{=}x'_j=x_i+x_j$ that $L(x'_{i\rightarrow x_i})=L((x'_{i\rightarrow x_i})_{j\rightarrow x_i})+L((x'_{i\rightarrow x_i})_{j\rightarrow x_j})\underset{((x'_{i\rightarrow x_i})_{j\rightarrow x_j})_i=x_i=((x'_{i\rightarrow x_i})_{j\rightarrow x_i})_j}{=}0+L((x'_{i\rightarrow x_i})_{j\rightarrow x_j})=L(x)$ so we have
$$L(x'_{i\rightarrow x_i})=L(x) \tag{10.63}$$
As $(x'_{i\rightarrow x_j})_j=x_i+x_j$ we have by multilinearity that $L(x'_{i\rightarrow x_j})=L((x'_{i\rightarrow x_j})_{j\rightarrow x_i})+L((x'_{i\rightarrow x_j})_{j\rightarrow x_j})\underset{\left((x'_{i\rightarrow x_j})_{j\rightarrow x_j}\right)_i=x_j=\left((x'_{i\rightarrow x_j})_{j\rightarrow x_j}\right)_j}{=}L((x'_{i\rightarrow x_j})_{j\rightarrow x_i})+0=L((x'_{i\rightarrow x_j})_{j\rightarrow x_i})$ giving
$$L(x'_{i\rightarrow x_j})=L((x'_{i\rightarrow x_j})_{j\rightarrow x_i}) \tag{10.64}$$

For every $k\in\{1,...,n\}$ we have $((x'_{i\rightarrow x_j})_{j\rightarrow x_j})_k=\begin{cases}(x'_{i\rightarrow x_j})_k \text{ if } k\in\{1,...,n\}\setminus\{j\}\\ x_i \text{ if } k=j\end{cases}=$ $\begin{cases}x'_k \text{ if } k\in\{1,...,n\}\setminus\{i,j\}\\ x_i \text{ if } k=j\\ x_j \text{if } k=i\end{cases}=\begin{cases}x_k \text{ if } k\in\{1,...,n\}\setminus\{i,j\}\\ x_i \text{ if } k=j\\ x_j \text{ if } k=i\end{cases}=(x\circ(i\leftrightarrow_m j))_k$ giving using 10.64 that

$$L(x'_{i\rightarrow x_j})=L(x\circ(i\leftrightarrow_n j)) \tag{10.65}$$

Using 10.62, 10.63 and 10.65 we have then $L(x)+L(x\circ(i\leftrightarrow_n j))\Rightarrow -L(x)=L(x\circ(i\leftrightarrow_n j))$ or

$$\text{If } i,j\in\{1,...,n\}, i\neq j \text{ then } (-1)\cdot L=(i\leftrightarrow_n j)L \tag{10.66}$$

We prove now by induction on m that if $\{(k_i\leftrightarrow l_i)\}_{i\in\{1,...,m\}}$ is a family where $k_i,l_i\in\{1,...,n\}$ and $k_i\neq l_i$ that

$$[(k_1\leftrightarrow_n l_1)\circ...\circ(k_m\leftrightarrow_n l_m)]L=(-1)^m\cdot L \tag{10.67}$$

Proof. So let $S=\{m\in\{1,...,\}|\text{if }\{(k_i\leftrightarrow_n l_i)\}_{i\in\{1,...,\}}$ is a family where k_i, $l_i\in\{1,...,n\}$ and $k_i\neq l_i$ then $[(k_1\leftrightarrow_n l_1)\circ...\circ(k_m\leftrightarrow l_m)]L=(-1)^m\cdot L\}$ then we have:

a. By 10.66 we have that $1\in S$

b. If $n\in S$ and if $\{(k_i\leftrightarrow l_i)\}_{i\in\{1,...,m+1\}}$ is a family where $k_i,l_i\in\{1,...,n\}$ and $k_i\neq l_i$ then $[(k_1\leftrightarrow_n l_1)\circ...\circ(k_{m+1}\leftrightarrow_n l_{m+1})]L=[[(k_1\leftrightarrow_n l_1)\circ...\circ(k_m\leftrightarrow_n l_m)]\circ(k_{m+1}\leftrightarrow l_{m+1})]\underset{10.83}{=}[(k_1\leftrightarrow_n l_1)\circ...\circ(k_m\leftrightarrow_n l_m)]((k_{m+1}\leftrightarrow_n l_{m+1})L)=(-1)\cdot[(k_1\leftrightarrow_n l_1)\circ...\circ(k_m\leftrightarrow_n l_m)]L\underset{m\in S}{=}(-1)\cdot((-1)^m\cdot L)\underset{(-1)^m=+/-1}{=}(-1)^{m+1}\cdot L$ proving that $m+1\in S$.

Using induction we have then that $S=\{1,...\}=\mathbb{N}$ and thus 10.67. □

If now $\sigma\in P_n$ then by 10.78 there exists a $\{(k_i\leftrightarrow_n l_i)\}_{i\in\{1,...,m\}}$ such that $k_i,l_i\in\{1,...,n\}$ and $k_i\neq l_i$ and $\sigma=(k_1\leftrightarrow_n l_1)\circ...\circ(k_m\leftrightarrow l_m)$ where by 10.89 we have that $(-1)^m=\mathrm{sign}(\sigma)$, using 10.67 we finally arrive at

$$\sigma L=\mathrm{sign}(\sigma)\cdot L$$

proving that L is skew symmetric.

3. **(2 ⇒ 3)** If $x=(x_1,...,x_n)\in X^n$ is such that $\{x_i\}_{i\in\{1,...,m\}}$ is linear dependent then using 10.144 there exists a $i\in\{1,...,n\}$, $\{\alpha_j\}_{j\in\{1,...,m\}\setminus\{i\}}$ such that $x_i=\sum_{j\in\{1,...,m\}\setminus\{i\}}\alpha_j\cdot x_j$ then as $x=x_{i\rightarrow x_i}$ we have $L(x)=L(x_{i\rightarrow x_i})\underset{\text{L is multilinear and 10.243}}{=}\sum_{j\in\{1,...,m\}\setminus\{i\}}\alpha_j\cdot L(x_1,...,x_{i-1},x_j,x_{i+1},...,x_n)\underset{j\neq i\,\text{and}\,(x_1,...,x_{i-1},x_j,x_{i+1},...,x_n)_j=x_j=(x_1,...,x_{i-1},x_j,...,x_n)_i}{=}\sum_{j\in\{1,...,m\}\setminus\{i\}}\alpha_j\cdot 0=0$

4. **(3 ⇒ 2)** If $x=(x_1,...,x_n)\in X^n$ is such that $\exists i,j\in\{1,...,n\}$ where $i\neq j$ and $x_i=x_j$ then $\{x_i\}_{i\in\{1,...,n\}}$ is linear dependent so we have $L(x)=0$ □

Corollary 10.258. *Let X,Y be vector spaces over a field F that has characteristics of zero then given $n\in\mathbb{N},n>1$, $L\in\mathrm{Hom}(X^n;Y)$ a skew symmetric n-linear mapping. If now X is finite dimensional with* $\dim(X)<n$ *then* $L=0$ *(the zero mapping)*

Proof. If $x=(x_1,\ldots,x_n)\in X^n$ then we have the following possibilities

1. $(\exists \boldsymbol{i}, \boldsymbol{j}\in\{1,\ldots,\boldsymbol{n}\}$ **with** $\boldsymbol{i}\neq \boldsymbol{j}$ **such that** $\boldsymbol{x_i}=\boldsymbol{x_j})$ then using the previous theorem we have as L is skew symmetric we have $L(x)$
2. $(\neg(\exists \boldsymbol{i}, \boldsymbol{j}\in\{1,\ldots,\boldsymbol{n}\}$ **with** $\boldsymbol{i}\neq \boldsymbol{j}$ **such that** $\boldsymbol{x_i}=\boldsymbol{x_j}))$ then we have $\forall i$, $j\in\{1,\ldots,n\}$ we have $x_i\neq x_j$ so $x:\{1,\ldots,n\}\rightarrow\{x_i|i\in\{1,\ldots,n\}\}$ is a bijection and thus $\#(x_i|i\in\{1,\ldots,n\})=n>\dim(X)$ so by 10.196 we have that $\{x_i|i\in\{1,\ldots,n\}\}$ is linear dependent and using 10.143 we have that $\{x_i\}_{i\in\{1,\ldots,n\}}$ is linear dependent which by the above theorem means that $L(x)=0$

So in all cases fro $x\in X^n$ we have $L(x)=0$ and thus $L=0$ □

Theorem 10.259. *Given $n\in\mathbb{N}$ X,Y vector spaces over a field F of characterization 0, $\dim(X)=n$ then if $\{e_i\}_{i\in\{1,\ldots,n\}}$ is a basis of X we have for $x=(x_1,\ldots,x_n)\in X^n$ that $\forall i\in\{1,\ldots,n\}$ there exists a $\{\xi_i^j\}_{j\in\{1,\ldots,n\}}$ such that $x_i=\sum_{j=1}^n \xi_i^j\cdot e_j$. Now if $L\in\mathrm{Hom}(X^n;Y)$ we have*

$$L(x_1,\ldots,x_n)=\sum_{i=(i_1,\ldots,i_n)\in\{1,\ldots,n\}^m}\left(\prod_{j\in\{1,\ldots,n\}}\xi_j^{i_j}\right)\cdot L(e_{i_1},\ldots,e_{i_n})$$

Proof. We proof this by induction on n so let $n\in\mathbb{N}$ and $S_n=\left\{m\in\{1,\ldots\}|\text{ if } m\leqslant n \text{ then } L(x_1,\ldots,x_n)=\sum_{(i_1,\ldots,i_m)\in\{1,\ldots,n\}^m}\left(\prod_{j\in\{1,\ldots,m\}}\xi_j^{i_j}\right)\cdot L(e_{i_1},\ldots e_{i_m},x_{m+1},\ldots,x_n)\right\}$ where $(e_{i_1},\ldots,e_{i_m},x_{m+1},\ldots,x_n)_k=\begin{cases} e_{i_k} \text{ if } k\in\{1,\ldots,m\}\\ x_k \text{ if } k\in\{m+1,\ldots,n\}\end{cases}$ we have then :

1. If $m=1$ then

$$\begin{aligned} L(x_1,\ldots,x_n) &= L\left(\sum_{i\in\{1,\ldots,n\}}\xi_1^i\cdot e_i,x_2,\ldots,x_n\right)\\ &= \sum_{i\in\{1,\ldots,n\}}\xi_1^i\cdot L(e_i,x_2,\ldots,x_n)\\ &= \sum_{i\{1,\ldots,n\}}\left(\prod_{j\in\{1\}}\xi_j^n\right)\cdot L(e_i,x_2,\ldots,x_m) \end{aligned}\tag{10.68}$$

Now $b:\{1,\ldots,n\}^1\rightarrow\{1,\ldots,n\}$ defined by $i\in\{1,\ldots,n\}^1\rightarrow b(i)=i_1$ is injective and bijective

a. **(injectivity)** if $b(i)=b(j)\Rightarrow\forall k\in\{1\}$ we have $i_k=i_1=b(i)=b(j)=j_1=j_k$ so $i=j$

b. **(surjectivity)** If $k\in\{1,\ldots,n\}$ then if we define $i\in\{1,\ldots,n\}^1$ by $i_1=k$ we have $b(i)=i_1=k$

Using 10.44 and the above equation we have

$$\begin{aligned} L(x_1,\ldots,x_n) &= \sum_{i=(i_1)\in\{1,\ldots,n\}^1}\left(\prod_{j\in\{1\}}\xi_j^{b(i_1)}\right)\cdot L(e_{b(i_1)},x_1,\ldots,x_n)\\ &\underset{b(i)=i_1}{=} \sum_{i=(i_1)\in\{1,\ldots,n\}^1}\left(\prod_{j\in\{1\}}\xi_j^{i_1}\right)\cdot L(e_{i_1},x_1,\ldots,x_n) \end{aligned}$$

proving that $1\in S_n$.

2. If $m\in S$ then if $m+1\leqslant n$ we have

$$\begin{aligned}
L(x_1,\ldots,x_n) &\underset{m\in S_n}{=} \sum_{i=(1,\ldots,n)^m}\left(\prod_{j\in\{1,\ldots,m\}}\xi_j^{i_j}\right)\cdot L(e_{i_1},\ldots,e_{i_m},x_{m+1},\ldots,x_n)\\
&= \sum_{i=(1,\ldots,n)^m}\left(\prod_{j\in\{1,\ldots,m\}}\xi_j^{i_j}\right)\cdot L\Bigg(e_{i_1},\ \ldots,\ e_{i_m},\\
&\qquad \sum_{k\in\{1,\ldots,n\}}\xi_{m+1}^k e_k,\ldots,x_n\Bigg)\\
&= \sum_{i\in\{1,\ldots,n\}^m}\Bigg(\sum_{k\in\{1,\ldots,n\}}\left(\left(\prod_{j\in\{1,\ldots,m\}}\xi_j^{i_j}\right)\cdot\xi_{m+1}^k\right)\cdot\\
&\qquad L(e_{i_1},\ldots,e_{i_m},e_k,\ldots,x_n)\Bigg)\\
&= \sum_{i\in\{1,\ldots,n\}^m}\left(\sum_{k\in\{1,\ldots m\}}A_{i,k}\right)
\end{aligned}$$

so we have

$$L(x_1,\ldots,x_n)=\sum_{i\in\{1,\ldots,n\}^m}\left(\sum_{k\in\{1,\ldots,m\}}A_{i,k}\right)\tag{10.69}$$

where

$$A_{i,k}=\left(\left(\prod_{j\in\{1,\ldots,m\}}\xi_j^{i_j}\right)\cdot\xi_{m+1}^k\right)\cdot L(e_{i_1},\ldots,e_{i_m},e_k,\ldots,x_n)\tag{10.70}$$

Given $i\in\{1,\ldots,n\}^m$ define now $I_i=\{i\}\times\{1,\ldots,n\}$ and b_i: $\{i\}\times\{1,\ldots,n\}\rightarrow\{1,\ldots,n\}$ by $(i,k)\rightarrow k$ which is trivially a bijection. Using 10.44 and 10.69 we have then $\sum_{k\in\{1,\ldots,m\}}A_{i,k}=\sum_{k\in I_i}A_{i,b_i(k)}\underset{k=(i,l)\Rightarrow k_1=i,b_i(k)=k_2}{=}\sum_{k\in I_i}A_{k_1,k_2}$

$$L(x_1,\ldots,x_n)=\sum_{i\in\{1,\ldots,n\}^m}\left(\sum_{k\in I_i}A_{k_1,k_2}\right)\tag{10.71}$$

If now $i,j\in\{1,\ldots,n\}^m$ and $i\neq j$ then if $m\in I_i\bigcap I_j$ we have that there exists a $k,l\in\{1,\ldots,n\}$ such that $m=(i,k)$ and $m=(j,l)$ but then $(i,k)=(j,l)\Rightarrow i=j$ contradicting $i\neq j$ so we have $I_i\bigcap I_j$. If we take now

$$I=\bigcup_{i\in\{1,\ldots,n\}^m}I_i\tag{10.72}$$

which is a finite union (see 10.104) of disjoint set, we can use the 10.47 to get

$$\sum_{i\in I} A_{i_1,i_2} = \sum_{i\in\{1,\ldots,n\}^m}\left(\sum_{k\in I_i} A_{k_1,k_2}\right) \tag{10.73}$$

Define now $h: \{1,\ldots,n\}^{m+1}\rightarrow\bigcup_{i\in\{1,\ldots,n\}^m} I_i$ by $i\rightarrow h(i) = (i_{|\{1,\ldots,m\}}, i_{m+1})$ then we have:

a. **(injectivity)** If $h(i) = h(j)$ then $(i_{|\{1,\ldots,m\}}, i_{m+1}) = (j_{|\{1,\ldots,m\}}, j_{m+1}) \Rightarrow i_{|\{1,\ldots,m\}} = j_{|\{1,\ldots,m\}}$ and $i_{m+1} = j_{m+1} \Rightarrow i = j$.

b. **(surjectivity)** If $z\in\bigcup_{i\in\{1,\ldots,n\}^m} I_i$ then there exists a $i\in\{1,\ldots, n\}^m$ such that $z\in I_i$ and thus there exists a $k\in\{1,\ldots,n\}$ such that $z = (i,k)$. Define then $l\in\{1,\ldots,n\}^{m+1}$ by $\forall r\in\{1,\ldots,n\}$, $l_r = \begin{cases} i_r \text{ if } r\in\{1,\ldots,m\} \\ k \text{ if } r = m+1\end{cases}$ then $h(l) = (l_{|\{1,\ldots,m}, l_{m+1}) = (i,k)$

which proves that h is a bijection. Using 10.44 and 10.73 we have then that

$$\sum_{i\in\{1,\ldots,n\}^{m+1}} A_{h(i)_1,h(i)_2} = \sum_{i\in I} A_{i_1,i_2} \tag{10.74}$$

Now $A_{h(i)_1,h(i)_2} = A_{i_{|\{1,\ldots,m\}},i_{m+1}} = \left(\left(\prod_{j\in\{1,\ldots,m\}}\xi_j^{(i_{|\{1,\ldots,m\}})_j}\right)\cdot\xi_{m+1}^{i_{m+1}}\right)\cdot L(e_{i_1},\ldots,e_{i_m},e_{i_{m+1}},\ldots e_n) = \left(\prod_{j\in\{1,\ldots,m+1\}}\xi_j^{i_j}\right)\cdot L(e_{i_1},\ldots,e_{i_{m+1}},\ldots,e_n)$ so that by using 10.70, 10.73 and the above gives finally that

$$L(x_1,\ldots,x_n) = \sum_{i\in\{1,\ldots,n\}^{m+1}}\left(\prod_{j\in\{1,\ldots,m+1\}}\xi_j^{i_j}\right)\cdot L(e_{i_1},\ldots,e_{i_{m+1}},\ldots,x_n)$$

proving that $m+1\in S_n$.

Using mathematical induction we have then $S_n = \{1,\ldots,\} = \mathbb{N}$ and thus if $n\in\mathbb{N} = S_n$ and $n\leqslant n$ we have $L(x_1,\ldots,x_n) = \sum_{i\in\{1,\ldots,n\}^m}\left(\prod_{j\in\{1,\ldots,n\}}\xi_j^{i_j}\right)\cdot L(e_{i_1},\ldots,e_{i_n})$ □

Theorem 10.260. *Given $n\in\mathbb{N}$ X,Y vector spaces over a field F of characterization 0, $\dim(X) = n$ then if $\{e_i\}_{i\in\{1,\ldots,n\}}$ is a basis of X we have for $x = (x_1,\ldots, x_n)\in X^n$ that $\forall i\in\{1,\ldots,n\}$ there exists a $\{\xi_i^j\}_{j\in\{1,\ldots,n\}}$ such that $x_i = \sum_{j=1}^n \xi_i^j\cdot e_j$. Now if $L\in\operatorname{Hom}(X^n;Y)$ is a skew symmetric n-linear mapping then*

$$L(x_1,\ldots,x_n) = \sum_{\sigma\in P_n}\left(\prod_{j\in\{1,\ldots,m\}}\xi_j^{\sigma(j)}\right)\cdot\sigma L(e_1,\ldots,e_n)$$

Proof. Now as P_n is the set of all bijections between $\{1,\ldots,n\}$ and $\{1,\ldots,n\}$ we have that $P_n\subseteq\{1,\ldots,n\}^n$. Now if $i\in\{1,\ldots,n\}^n\setminus P_n$ then i is not a bijection and thus not a injection [if i would be injective then $i: \{1,\ldots,n\}\rightarrow i(\{1,\ldots, n\})$ would be a bijection, and if $i(\{1,\ldots,n\}) = \{1,\ldots,n\}$ we would have that i is a bijection so we must have that $i(\{1,\ldots,n\})\subset\{1,\ldots,n\}$ giving $n = \#(\{1,\ldots, n\}) \underset{\#\{1,\ldots,n\}\approx\#(\{1,\ldots,n\})}{=} \#(i((1,\ldots,n))) < \#(\{1,\ldots,n\}) = n$ reaching the contradiction $n<n$], so there exists a $k,l\in\{1,\ldots,n\}$ such that $i_k=i_l\Rightarrow e_{i_k}=e_{i_l}$ and thus $L(e_{i_1},\ldots,e_{i_n}) \underset{L\,\text{is skew symmetric and } 10.257}{=} 0$. Using the previous theorem we have then

$$\begin{aligned}
L(x_1,\ldots,x_n) &= \sum_{\sigma\in\{1,\ldots,n\}^n}\left(\prod_{j\in\{1,\ldots,n\}}\xi_j^{\sigma j}\right)\cdot L(e_{\sigma_1},\ldots,e_{\sigma_n})\\
&= \sum_{\sigma\in\{1,\ldots,n\}^n\setminus P_n}\left(\prod_{j\in\{1,\ldots,n\}}\xi_j^{\sigma j}\right)\cdot L(e_{\sigma_1},\ldots,e_{\sigma_n})+\\
&\quad \sum_{\sigma\in P_n}\left(\prod_{j\in\{1,\ldots,n\}}\xi_j^{\sigma j}\right)\cdot L(e_{\sigma_1},\ldots,e_{\sigma_n})\\
&= \sum_{\sigma\in\{1,\ldots,n\}^n\setminus P_n}\left(\prod_{j\in\{1,\ldots,n\}}\xi_j^{\sigma j}\right)\cdot 0+\sum_{\sigma\in P_n}\left(\prod_{j\in\{1,\ldots,n\}}\xi_j^{\sigma(j)}\right)\cdot\\
&\quad L(e_{\sigma_{(1)}},\ldots,e_{\sigma(n)})\\
&= \sum_{\sigma\in P_n}\left(\prod_{j\in\{1,\ldots,n\}}\xi_j^{\sigma(j)}\right)\cdot L(e_{\sigma_{(1)}},\ldots,e_{\sigma(n)})\\
&= \sum_{\sigma\in P_n}\left(\prod_{j\in\{1,\ldots,n\}}\xi_j^{\sigma(j)}\right)\cdot \sigma L(e_1,\ldots,e_n)
\end{aligned}$$

□

Corollary 10.261. *Given $n\in\mathbb{N}$ X,Y vector spaces over a field F of characterization 0, $\dim(X)=n$ then if $\{e_i\}_{i\in\{1,\ldots,n\}}$ is a basis of X we have for $x=(x_1,\ldots,x_n)\in X^n$ that $\forall i\in\{1,\ldots,n\}$ there exists a $\{\xi_i^j\}_{j\in\{1,\ldots,n\}}$ such that $x_i=\sum_{j=1}^n\xi_i^j\cdot e_j$. Now if $L\in\operatorname{Hom}(X^n;Y)$ is a skew symmetric n-linear mapping then*

$$L(x_1,\ldots,x_n)=\left(\sum_{\sigma\in P_n}\operatorname{sign}(\sigma)\cdot\prod_{j\in\{1,\ldots,m\}}\right)\cdot L(e_1,\ldots,e_n)$$

Proving that L depends only on the values of L at $(e_1,\ldots,e_n)$ [in other words if L_1, L_2 are two skew symmetric n-linear mappings such that $L_1(e_1,\ldots,e_n)=L_2(e_1,\ldots,e_n)$ then $L_1=L_2$]. This also proves that if L is a n-linear skew symmetric mapping such that $L(e_1,\ldots,e_n)=0$ then $L=0$.

Proof. As L is a n-linear skew symmetric mapping we have using the previous theorem that

$$\begin{aligned}
L(x_1,\ldots,x_n) &= \sum_{\sigma\in P_n}\left(\prod_{j\in\{1,\ldots,n\}}\xi_j^{\sigma(j)}\right)\cdot\sigma L(e_1,\ldots,e_n)\\
&\underset{L\,\text{is skew symmetric}}{=} \sum_{\sigma\in P_n}\left(\prod_{j\in\{1,\ldots,n\}}\xi_j^{\sigma(j)}\right)\cdot(\operatorname{sign}(\sigma)\cdot L(e_1,\ldots,e_n))\\
&= \left(\sum_{\sigma\in P_n}\operatorname{sign}(\sigma)\cdot\prod_{j\in\{1,\ldots,m\}}\right)\cdot L(e_1,\ldots,e_n)
\end{aligned}$$

□

Definition 10.262. *Given $n \in \mathbb{N}$ let X be a n-dimensional vector space over a field F then a determinant function Δ in X is a skew symmetric n-linear mapping $\Delta: X^n \rightarrow F$*

Theorem 10.263. *Given $n \in \mathbb{N}$ let X be a n-dimensional vector space over a field F of characterization zero then if $\{e_i\}_{i \in \{1,\ldots,n\}}$ is a basis for X there exists a determinant function Δ such that $\Delta(e_1, \ldots, e_n) = 1$ (the unit in F). This proves that there exist a non trivial determinant function (meaning a determinant function different from the zero mapping).*

Proof. Let $\{e_i\}_{i \in \{1,\ldots,n\}}$ be a basis in X then $\forall i \in \{1, \ldots, n\}$ define $f_i: X \rightarrow F$ by $x = \sum_{i \in 1,\ldots,n} \alpha_i \cdot x_i \rightarrow f_i(x) = \alpha_i$ (which is well defined as for every $x \in X$ there exist a unique expansion of x in the basis e_i) and is linear [as for if $\alpha, \beta \in F$ and x, $y \in X$ with $x = \sum_{i \in \{1,\ldots,n\}} \alpha_i \cdot e_i$, $y = \sum_{i \in \{1,\ldots,n\}} \beta_i \cdot e_i$ then $f_i(\alpha \cdot x + \beta \cdot y) = f_i\big(\alpha \cdot \sum_{j \in \{1,\ldots,n\}} \alpha_j \cdot e_j + \beta \cdot \sum_{j \in \{1,\ldots,n\}} \beta_j \cdot e_j\big) = f_i\big(\sum_{j \in \{1,\ldots,n\}} (\alpha \cdot \alpha_j + \beta \cdot \beta_j) \cdot e_j\big) = \alpha \cdot \alpha_i + \beta \cdot \beta_i = \alpha \cdot f_i(x) + \beta \cdot f_i(y)$]. As $\forall j \in \{1, \ldots, n\}$ we have $e_j = \sum_{i \in \{1,\ldots,n\}} \delta_{i,j} \cdot e_i$ we have that $f_i(e_j) = \delta_{i,j}$. Given $x = (x_1, \ldots, x_n) \in X^n$ define then

$$\Phi: X^n \rightarrow F \text{ by } (x_1, \ldots, x_n) \rightarrow \Phi(x_1, \ldots, x_n) = \prod_{j \in \{1,\ldots,n\}} f_j(x_j)$$

If $i \in \{1, \ldots, n\}$, $\alpha, \beta \in F$ and $x, y \in X$ then we have

$$\begin{aligned}
\Phi(x_1, \ldots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_n) &= f_i(\alpha \cdot x + \beta \cdot y) \cdot \prod_{j \in \{1,\ldots,n\} \setminus \{i\}} f_j(x_j) \\
&= (\alpha \cdot f_i(x) + \beta \cdot f_i(y)) \cdot \prod_{j \in \{1,\ldots,n\} \setminus \{i\}} f_j(x_j) \\
&= \alpha \cdot \left(f_i(x) \cdot \prod_{j \in \{1,\ldots,n\} \setminus \{i\}} f_j(x_j) \right) + \beta \cdot \\
&\quad \left(f_i(y) \cdot \prod_{j \in \{1,\ldots,n\} \langle \setminus \{i\} \rangle} f_j(x_j) \right) \\
&= \alpha \cdot \Phi(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) + \beta \cdot \\
&\quad \Phi(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)
\end{aligned}$$

proving that Φ is n-linear. Using 10.256 we create the skew symmetric n-linear mapping

$$\Delta: X^n \rightarrow F \text{ by } \Delta = \sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \sigma\Phi$$

that is then by definition a determinant function. Now as $f_i(e_j) = \delta_{i,j}$ we have if $\sigma \in P_n \setminus 1_{\{1,\ldots,n\}}$ that there exists a $i \in \{1, \ldots, n\}$ such that $\sigma(i) \neq i$ and thus $f_i(e_{\sigma(i)}) = 0$ and thus $\sigma\Phi(e_1, \ldots, e_n) = \prod_{j \in \{1,\ldots,n\}} f_j(e_{\sigma(j)}) = f_i(e_i) \cdot \prod_{j \in \{1,\ldots,n\} \setminus \{i\}} f_j(e_j) = 0 \cdot \prod_{j \in \{1,\ldots,n\} \setminus \{i\}} f_j(e_j) = 0$, if $\sigma = 1_{\{1,\ldots,n\}}$ then $\sigma\Phi(e_1, \ldots, e_n) = \prod_{j \in \{1,\ldots,n\}} f_j(e_j) = \prod_{j \in \{1,\ldots,m\}} 1 = 1$ this gives then

$$\sigma\Phi(e_1, \ldots, e_n) = \begin{cases} 0 \text{ if } \sigma \neq 1_{\{1,\ldots,n\}} \\ 1 \text{ if } \sigma = 1_{\{1,\ldots,n\}} \end{cases}$$

So we have finally $\Delta(e_1, \ldots, e_n) = \sum_{\sigma \in P_n} \mathrm{sign}(\sigma) \cdot \sigma\Phi(e_1, \ldots, e_n) = \sum_{\sigma \in \{i_{\{1,\ldots,n\}}\}} \mathrm{sign}(\sigma) \cdot \sigma\Phi(e_1,\ldots,e_n) + \sum_{\sigma \in P_n \setminus \{i_{\{1,\ldots,n\}}\}} \mathrm{sign}(\sigma) \cdot \sigma\Phi(e_1,\ldots,e_n) = 1 + 9 = 1$

□

Theorem 10.264. *Given $n \in \mathbb{N}$ let X be a n-dimensional vector space over a field F of characterization zero, Y a vector space over F and Δ a non zero determinant function. Then for every skew symmetric n-linear function $L: X^n \to Y$ there exists a unique $y \in Y$ such that $\forall (x_1, \ldots, x_n) \in X^n$ we have that $L(x_1, \ldots, x_n) = \Delta(x_1, \ldots, x_n) \cdot y$*

Proof. Let $\{e_i\}_{i \in \{1,\ldots,n\}}$ a basis of X then using 10.261 and the fact that $\Delta \neq 0$ we must have that $\Delta(e_1, \ldots, e_n) \neq 0$. Define now $\{e'_i\}_{i \in \{1,\ldots,n\}}$ by $e'_i = \begin{cases} \Delta(e_1, \ldots, e_n)^{-1} \cdot e_1 \text{ if } i = 1 \\ e_i \text{ if } i \in \{2, \ldots, n\} \end{cases}$. We prove now that $\{e'_i\}_{i \in \{1,\ldots,n\}}$ is still a basis for X.

Proof. If $\{\alpha_i\}_{i \in \{1,\ldots,n\}}$ is a a family in F such that $\sum_{i \in \{1,\ldots,n\}} \alpha_i \cdot e'_i = 0$ then if we define $\{\alpha'_i\}_{i \in \{1\ldots,n\}}$ by $\alpha'_i = \begin{cases} \Delta(e_1, \ldots, e_n)^{-1} \cdot \alpha_1 \text{ if } i = 1 \\ \alpha_i \text{ if } i \in \{2, \ldots, n\} \end{cases}$ we have $\alpha'_i \cdot e_i = \begin{cases} \alpha_1 \cdot \Delta(e_1, \ldots, e_n)^{-1} \cdot e_1 = \alpha_1 \cdot e'_1 \text{ if } i = 1 \\ \alpha_i \cdot e'_i \text{ if } i \in \{2, \ldots, n\} \end{cases} = \alpha_i \cdot e'_i$ and $\sum_{i \in \{1,\ldots,n\}} \alpha'_i \cdot e_i = \sum_{i \in \{1,\ldots,n\}} \alpha_i \cdot e'_i = 0$. So $\forall i \in \{1, \ldots, n\}$we have $\alpha'_i = 0 \Rightarrow 0 = \begin{cases} \Delta(e_1, \ldots, e_n)^{-1} \cdot \alpha_1 \text{ if } i = 1 \\ \alpha_i \text{ if } i \in \{2, \ldots n\} \end{cases} \Rightarrow \alpha_i = 0$ proving linear Independence. If $x \in X$ then there exists a $\{\beta_i\}_{i \in \{1,\ldots,n\}}$ such that $\sum_{i \in \{1,\ldots,n\}} \beta_i \cdot e_i = x$, so if we take $\{\beta'_i\}_{i \in \{1,\ldots,n\}}$ where $\beta'_i = \begin{cases} \Delta(e_1, \ldots, e_n) \cdot \beta \text{ if } i = 1 \\ \beta_i \text{ if } i \in \{2, \ldots, n\} \end{cases}$ then we have $\beta'_i \cdot e'_i = \begin{cases} \Delta(e_1, \ldots, e_n) \cdot (\Delta(e_1, \ldots, e_n)^{-1} \cdot e_1) \text{ if } i = 1 \\ \beta_i \cdot e_i \text{ if } i \in \{2, \ldots, n\} \end{cases} = \beta_i \cdot e_i$. So $\sum_{i \in \{1,\ldots,n\}} \beta'_i \cdot e'_i = \sum_{i \in \{1,\ldots,n\}} \beta_i \cdot e_i = x$ □

Now $\Delta(e'_1, \ldots, e'_n) = \Delta(\Delta(e_1, \ldots, e_n)^{-1} \cdot e_1, \ldots, e_n) = \Delta(e_1, \ldots, e_n)^{-1} \cdot \Delta(e_1, \ldots, e_n) = 1$. Set $y = L(e'_1, \ldots, e'_n)$ and define $K: X^n \to Y$ by $(x_1, \ldots, x_n) \to K(x_1, \ldots, x_n) = \Delta(x_1, \ldots, x_n) \cdot y$, then because Δ is a skew symmetric n-linear mapping it is trivial to show that K is also skew symmetric and n-linear mapping. Now $K(e'_1, \ldots, e'_n) = \Delta(e'_1, \ldots, e'_n) \cdot y = 1 \cdot y = y = L(e'_1, \ldots, e'_n)$ and using 10.261 we have then $K = L$ proving that $L(x_1, \ldots, x_n) = \Delta(x_1, \ldots, x_n) \cdot y$. Now to prove that y is unique assume that also $L(x_1, \ldots, x_n) = \Delta(x_1, \ldots, x_n) \cdot y'$ then $\Delta(e'_1, \ldots, e'_n) \cdot y = \Delta(e'_1, \ldots, e'_n) \cdot y' \Rightarrow 1.y = 1.y' \Rightarrow y = y'$. □

Corollary 10.265. *Given $n \in \mathbb{N}$ let X be a n-dimensional vector space over a field F of characterization zero. Let Δ be a non zero determinant function then every other determinant function is a scalar multiple of Δ*

Proof. If $\Delta': X^n \to F$ is a another determinant function then by the above theorem there exists a $y \in F$ such that $\Delta' = \Delta \cdot y$ □

Definition 10.266. *Let $n \in \mathbb{N}$, X a set then given $i \in \{1, \ldots, n\}$ and $y \in X$ and $x = (x_1, \ldots, x_n) \in X^n$ we define $(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in X^n$ to be defined by $\forall k \in \{1, \ldots, n\}$ we have $(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k = \begin{cases} y \text{ if } k = 1 \\ x_{k-1} \text{ if } k \in \{2, \ldots, i\} \\ x_k \text{ if } k \in \{i+1, \ldots, n\} \end{cases}$*

Example 10.267. if $n=1$ and $x=(x_1)\in X^1$ and $y\in X$ then by the above definition we have that $(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)=(y)$

Lemma 10.268. *Let $n\in\mathbb{N}$, X a vector space over a field F then given $i\in\{1,\ldots,n\}$, $j\in\{1,\ldots,n\}$ and $y\in X$ and $x=(x_1,\ldots,x_n)\in X^n$ then we have*

1. $\boldsymbol{(j+1\leqslant i)}$ *then* $((y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{j+1\rightarrow x_i}\circ(j+1\rightsquigarrow_n i))=(y,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)$
2. $\boldsymbol{(i+1\leqslant j)}$ *then* $((y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{j\rightarrow x_i}\circ(j\rightsquigarrow_n i+1))=(y,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)$

Proof.

1. $\boldsymbol{(j+1<i)}$ Let $k\in\{1,\ldots,n\}$ then we have (using the definition of $\rightsquigarrow$) the following cases to consider for $A_k=((y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{j+1\rightarrow x_i}\circ(j+1\rightsquigarrow_n i))_k$:

   a. $\boldsymbol{(k=1)}$ then

$$\begin{array}{rcl} A_k & \underset{k=1<j+1}{=} & ((y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{j+1\rightarrow x_i})_k \\ & \underset{k=1\neq j+1}{=} & (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k \\ & \underset{k=1}{=} & y \\ & \underset{k=1}{=} & (y,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k \end{array}$$

   b. $\boldsymbol{(k\neq 1\wedge k<j+1)}$ then

$$\begin{array}{rcl} A_k & = & ((y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{j+1\rightarrow x_i})_k \\ & \underset{k\neq j+1}{=} & (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k \\ & \underset{k\neq 1\wedge k<j+1\leqslant i}{=} & x_{k-1} \\ & \underset{k\neq 1\wedge k<j+1\Rightarrow k\leqslant j}{=} & (y,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k \end{array}$$

   c. $\boldsymbol{(k\neq 1\wedge j+1\leqslant k<i)}$ then

$$\begin{array}{rcl} A_k & = & ((y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{j+1\rightarrow x_i})_{k+1} \\ & \underset{j+1\leqslant k<k+1}{=} & (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{k+1} \\ & \underset{k\neq 1\wedge k<i+1\Rightarrow k+1\leqslant i}{=} & (y,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k \end{array}$$

   d. $\boldsymbol{(k\neq 1\wedge k=i)}$ then

$$\begin{array}{rcl} A_k & = & ((y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{j+1\rightarrow x_i})_{j+1} \\ & = & x_i \\ & \underset{j<i=k}{=} & (y,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k \end{array}$$

e. $(\boldsymbol{k \neq 1 \wedge i < k})$ then

$$\begin{aligned} A_k &= ((y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j+1 \to x_i})_k \\ &\underset{j+1<i<k}{=} (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\ &= x_k \\ &\underset{j+1<i<k}{=} (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \end{aligned}$$

So $\forall k \in \{1, \ldots, n\}$ we have $((y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j+1 \to x_i} \circ (j+1 \rightsquigarrow_n i))_k = (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k$ or $(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j+1 \to x_i} \circ (j + 1 \rightsquigarrow_n i) = (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$.

2. $(\boldsymbol{j+1 = i})$ Then by the definition of $(j+1 \rightsquigarrow_n i)$ we have $(j+1 \rightsquigarrow_n i) = 1_{\{1, \ldots, n\}}$ and thus $(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j+1 \to x_i} \circ (j+1 \rightsquigarrow_n i) = (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j+1 \to x_i}$ and we have then $\forall k \in \{1, \ldots, n\}$ to consider for $A_k = (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j+1 \to x_i} \circ (j + 1 \rightsquigarrow_n i)_k = ((y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j+1 \to x_i})_k$ the following cases:

a. $(\boldsymbol{k = 1})$ then

$$\begin{aligned} A_k &\underset{k=1<j+1}{=} (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\ &\underset{k=1}{=} y \\ &\underset{k=1}{=} (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \end{aligned}$$

b. $(\boldsymbol{k \neq 1 \wedge k < j})$ then

$$\begin{aligned} A_k &\underset{k<j<j+1}{=} (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\ &\underset{k<j<j+1=i}{=} x_{k-1} \\ &\underset{k\neq 1 \wedge k<j}{=} (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \end{aligned}$$

c. $(\boldsymbol{k \neq 1 \wedge k = j})$ then

$$\begin{aligned} A_k &\underset{k=j\neq j+1}{=} (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\ k = j < j+1 = i\ & x_{k-1} \\ &\underset{k\neq 1 \wedge k=j\leqslant j}{=} (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k \end{aligned}$$

d. $(\boldsymbol{k \neq 1 \wedge k = j+1})$ then

$$\begin{aligned} A_k &\underset{k=j+1}{=} x_i \\ &\underset{k=j+1=i}{=} x_k \\ &\underset{j<j+1=k}{=} (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \end{aligned}$$

e. $(\boldsymbol{k \neq 1 \wedge j+1 < k})$ then

$$\begin{aligned} A_k &\underset{j+1<k}{=} (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\ &\underset{i=j+1<k \Rightarrow i+1\leqslant k}{=} x_k \\ &\underset{j<j+1<k}{=} (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \end{aligned}$$

So $\forall k \in \{1, \ldots, n\}$ we have $((y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j+1 \rightarrow x_i} \circ (j+1 \rightsquigarrow_n i))_k = (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k$ or $(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j+1 \rightarrow x_i} \circ (j + 1 \rightsquigarrow_n i) = (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$.

3. $(\boldsymbol{i+1<j})$ Let $k \in \{1, \ldots, n\}$ then using the definition of $\rightsquigarrow$ we have to consider the following cases for $A_k = ((y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j \rightarrow x_i} \circ (j \rightsquigarrow_n i + 1))_k$

   a. $(k=1)$ then

$$\begin{array}{lcl} A_k & \underset{k=1<i+1}{=} & ((y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j \rightarrow x_i})_k \\ & \underset{k=1<i+1<j}{=} & (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\ & \underset{k=1}{=} & y \\ & \underset{k=1}{=} & (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \end{array}$$

   b. $(\boldsymbol{k \neq 1 \wedge k < i+1})$ then

$$\begin{array}{lcl} A_k & = & ((y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j \rightarrow x_i})_k \\ & \underset{k<i+1<j}{=} & (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\ & \underset{k \neq 1 \wedge k<i+1 \Rightarrow k \leqslant i}{=} & x_{k-1} \\ & \underset{k<i+1<j}{=} & (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \end{array}$$

   c. $(\boldsymbol{k \neq 1 \wedge k = i+1})$ then

$$\begin{array}{lcl} A_k & = & ((y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j \rightarrow x_i})_j \\ & = & x_i \\ & \underset{k \neq 1 \wedge k=i+1<j}{=} & (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{k=i+1} \end{array}$$

   d. $(\boldsymbol{k \neq 1 \wedge i+1 < k \leqslant j})$ then

$$\begin{array}{lcl} A_k & = & ((y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j \rightarrow x_i})_{k-1} \\ & \underset{k \leqslant j \Rightarrow k-1<j}{=} & (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{k-1} \\ & \underset{i+1<k \Rightarrow i<k-1}{=} & x_{k-1} \\ & \underset{k \neq 1 \wedge k \leqslant j}{=} & (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \end{array}$$

   e. $(\boldsymbol{k \neq 1 \wedge j < k})$ then

$$\begin{array}{lcl} A_k & = & ((y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j \rightarrow x_i})_k \\ & \underset{j<k}{=} & (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\ & \underset{i+1<j<k \Rightarrow i<k}{=} & x_k \\ & \underset{k \neq 1 \wedge j<k}{=} & (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \end{array}$$

So $\forall k \in \{1, \ldots, n\}$ we have $((y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j \to x_i} \circ (j \rightsquigarrow_n i+1))_k = (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k$ and thus $(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j \to x_i} \circ (j \rightsquigarrow_n i+1) = (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$

4. $(\boldsymbol{i+1=j})$ Then by the definition of $(j \rightsquigarrow_n i+1)$ we have that $(j \rightsquigarrow_n i+1) = 1_{\{1, \ldots, n\}}$ and thus we have that $(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j \to x_i} \circ (j \rightsquigarrow_n i+1) = (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j \to x_i}$ then $\forall k \in \{1, \ldots, n\}$ we have for $A_k = ((y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j \to x_i} \circ (j \rightsquigarrow_n i+1))_k = ((y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j \to x_i})_k$ the following cases to consider

   a. $(\boldsymbol{k=1})$ then

$$\begin{aligned} A_k &\underset{k=1\neq i+1=j}{=} (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\ &\underset{k=1}{=} y \\ &\underset{k=1}{=} (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \end{aligned}$$

   b. $(\boldsymbol{k \neq 1 \wedge k < i})$ then

$$\begin{aligned} A_k &\underset{k<i<i+1=j}{=} (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\ &\underset{k\neq 1 \wedge k<i}{=} x_{k-1} \\ &\underset{k\neq 1\wedge k<i<i+1=j}{=} (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \end{aligned}$$

   c. $(\boldsymbol{k \neq 1 \wedge k = i})$ then

$$\begin{aligned} A_k &\underset{k=i\neq i+1=j}{=} (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\ &\underset{k=i\leqslant i}{=} x_{k-1} \\ &\underset{k=i<i+1=j}{=} (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \end{aligned}$$

   d. $(\boldsymbol{k \neq 1 \wedge k = i+1})$ then

$$\begin{aligned} A_k &\underset{k=i+1=j}{=} x_i \\ &\underset{k=i+1}{=} x_{k-1} \\ &\underset{k=i+1=j\leqslant j}{=} (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \end{aligned}$$

   e. $(\boldsymbol{k \neq 1 \wedge i+1 < k})$ then

$$\begin{aligned} A_k &\underset{j=i+1<k}{=} (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\ &\underset{i<i+1<k}{=} x_k \\ &\underset{j=i+1<k}{=} (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \end{aligned}$$

So $\forall k \in \{1, \ldots, n\}$ we have $((y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j \to x_i} \circ (j \rightsquigarrow_n i+1))_k = (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k$ and thus $(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j \to x_i} \circ (j \rightsquigarrow_n i+1) = (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$

□

Definition 10.269. *Let $n \in \mathbb{N}$ then if X is a vector space over a field F and $\Delta: X^n \to F$ a determinant function in X then we define $\hat{\Delta}: X \times X^n \to X$ by $(y, (x_1, \ldots, x_n)) \to \hat{\Delta}(y, (x_1, \ldots, x_n)) = \sum_{j \in \{1, \ldots, n\}} (-1)^{j-1} \Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j$*

Theorem 10.270. *Let $n \in \mathbb{N}$ then if X is a n-dimensional vector space over a field F of characterization zero and $\Delta: X^n \to F$ a determinant function in X if $y \in X$ and $(x_1, \ldots, x_n) \in X^n$ then $\hat{\Delta}(y, (x_1, \ldots, x_n)) = \Delta(x_1, \ldots, x_n) \cdot y$*

Proof. Then we have to consider the following cases for $(x_1, \ldots, x_n) \in X^n$:

1. **($\{x_i\}_{i \in \{1, \ldots, n\}}$ are linear dependent)** then by 10.257 we have that $\Delta(x_1, \ldots, x_n) = 0$ we must prove that $\hat{\Delta}(y, (x_1, \ldots, x_n)) = 0$. Consider the following cases for n:

   a. **($n = 1$)** In this case let $\{e_1\}$ be the basis of X then if $y \in X$ and $x = (x_1) \in X^1$ there exists a $\alpha, \beta \in F$ such that $y = \alpha \cdot e_1$, $x_1 = \beta \cdot e_1$ and thus $\hat{\Delta}(y, (x_1)) \underset{10.267}{=} (-1)^0 \cdot \Delta(y) \cdot x_1 = \Delta(\alpha \cdot e_1) \cdot (\beta \cdot e_1) = \Delta(\beta \cdot e_1) \cdot (\alpha \cdot e_1) = \Delta(x_1) \cdot y$

   b. **($n = 2$)** In this case $\hat{\Delta}(y, (x_1, x_2)) = (-1)^0 \cdot \Delta(y, x_2) \cdot x_1 + (-1)^1 \cdot \Delta(y, x_1) \cdot x_2 = \Delta(y, x_2) \cdot x_1 - \Delta(u, x_1) \cdot x_2$. Now as $\{x_i\}_{i \in \{1, \ldots, 2\}}$ is linear dependent there exists by 10.144 a $\alpha \neq 0$ such that $x_1 = \alpha \cdot x_2$ and thus $\hat{\Delta}(y, (x_1, x_2)) = \Delta(y, x_2) \cdot (\alpha \cdot x_2) - \Delta(y, \alpha \cdot x_2) \cdot x_2 = \alpha \cdot (\Delta(y, x_2) \cdot x_2) - \alpha \cdot (\Delta(y, x_2) \cdot x_2) = 0$

   c. **($n > 2$)** In this case as $\{x_i\}_{i \in \{1, \ldots, n\}}$ are linear independent there exists by 10.144 a $i_0 \in \{1, \ldots, n\}$ and $\{\alpha_i\}_{i \in \{1, \ldots, n\} \setminus \{i_0\}}$ such that

$$x_{i_0} = \sum_{i \in \{1, \ldots, n\} \setminus \{i_0\}} \alpha_i \cdot x_i \tag{10.75}$$

   Now $\hat{\Delta}(y, (x_1, \ldots, x_n)) = \sum_{j \in \{i_0\}} (-1)^j \cdot \Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j + \sum_{j \in \{1, \ldots, n\} \setminus \{i_0\}} (-1)^j \cdot \Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j = (-1)^{i_0} \cdot \Delta(y, x_1, \ldots x_{i_0-1}, x_{i_0+1}, \ldots, x_n) \cdot x_{i_0} + \sum_{j \in \{1, \ldots, n\} \setminus \{i_0\}} (-1)^j \cdot \Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j$ giving

$$\hat{\Delta}(y, (x_1, \ldots, x_n)) = A + B \tag{10.76}$$

where

$$A = (-1)^{i_0} \cdot \Delta(y, x_1, \ldots, x_{i_0-1}, x_{i_0+1}, \ldots, x_n) \cdot x_{i_0}$$

$$B = \sum_{j \in \{1,\ldots,n\} \setminus \{i_0\}} (-1)^j \cdot \Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j$$

Using 10.75 we have then that

$$A = (-1)^{j_0} \cdot \Delta(y, x_1, \ldots, x_{i_0-1}, x_{i_0+1}, \ldots, x_n) \cdot \sum_{i \in \{1,\ldots,n\} \setminus \{i_0\}} \alpha_i \cdot x_i \tag{10.77}$$

Now for $j \in \{1, \ldots, n\} \setminus \{i_0\}$ we have two cases to consider:

i. $\boldsymbol{(i_0 < j)}$ then using 10.75 and the multilinearity of Δ we have

$$\Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j = \sum_{k \in \{1,\ldots,n\} \setminus \{i_0\}} \alpha_k \cdot \Delta((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0+1 \rightarrow x_k}) \cdot x_j \tag{10.78}$$

now for $k \in \{1, \ldots, n\} \setminus \{i_0\}$ consider the following cases

A. $\boldsymbol{(k < j)}$ then $((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0+1 \rightarrow x_k})_{k+1} \underset{k+1 < j+1 < i_0+1}{=} (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{k+1} \underset{1 < k+1}{=} x_k$ and $((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0+1 \rightarrow x_k})_{i_0+1} = x_k$ this together with $k \neq i_0 \Rightarrow k + 1 \neq i_0 + 1$ means by 10.257 that

$$\Delta((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0+1 \rightarrow x_k}) = 0$$

B. $\boldsymbol{(j < k)}$ then $((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0+1 \rightarrow x_k})_k \underset{i_0 < j < k \Rightarrow i_0+1 \leqslant j < k}{=} (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \underset{j < k}{=} x_k$ and $((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0+1 \rightarrow x_k})_{i_0+1} = x_k$ this together with $i_0 < j < k \Rightarrow i_0 + 1 \neq k$ and 10.257 that

$$\Delta((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0+1 \rightarrow x_k}) = 0$$

Using (1) and (2) we conclude that $\Delta((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0+1 \rightarrow x_k}) = \delta_{j,k} \cdot \Delta((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0+1 \rightarrow x_k})$ and thus using 10.78 we have $\Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j = \sum_{k \in \{1,\ldots,n\} \setminus \{i_0\}} \delta_{j,k} \cdot (\alpha_k \cdot (\Delta((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0+1 \rightarrow x_k}) \cdot x_j)) = \alpha_j \cdot \Delta((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0+1 \rightarrow x_j}) \cdot x_j$ giving :

$$\Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j = \alpha_j \cdot (\Delta((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0+1 \rightarrow x_j}) \cdot x_j) \tag{10.79}$$

Now $(i_0+1 \rightsquigarrow j)\Delta((y,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{i_0+1\to x_j}) = \Delta((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0+1\to x_j}) \underset{i_0<j\Rightarrow i_0+1\leqslant j \text{ and } 10.268}{=} \Delta(y, x_1, \ldots, x_{i_0-1}, x_{i_0+1}, \ldots, x_n)$ using 10.96 and the fact that Δ is skew symmetric and $i_0+1<j$ we have then that $(-1)^{j-i_0-1}\Delta((y,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{i_0+1\to x_i}) = (i_0+1 \rightsquigarrow j)\Delta((y,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{i_0+1\to x_j}) = \Delta(y,x_1,\ldots,x_{i_0-1}, x_{i_0+1},\ldots,x_n)$ and by multiplying both sides by $(-1)^{j-i_0-1}$ and the fact that $1=(-1)^{j-i_0-1}\cdot(-1)^{j-i_0-1}$ gives using 10.79 that

$$\boldsymbol{\Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j = (-1)^{j-i_0-1} \cdot (\alpha_j \cdot (\Delta(y, x_1, \ldots, x_{i_0-1}, x_{i_0+1}, \ldots, x_n) \cdot x_j))} \tag{10.80}$$

ii. $\boldsymbol{(j<i_0)}$ Then using 10.75 and the multilinearity of Δ we have that

$$\Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j = \sum_{k\in\{1,\ldots,n\}\setminus\{i_0\}} \alpha_k \cdot \Delta((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0\to x_k}) \cdot x_j \tag{10.81}$$

now for $k\in\{1,\ldots,n\}\setminus\{i_0\}$ consider the following cases:

A. $\boldsymbol{(k<j)}$ then $((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0\to x_k})_{k+1} \underset{k<j<i_0\Rightarrow k+1\leqslant k<i_0}{=} (y,x_1,\ldots,x_{j-1},x_{j+1},\ldots, x_n)_{k+1} = x_k$ and $((y,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{i_0\to x_k})_{i_0} = x_k$, this together with $k+1\leqslant j<i_0$ and 10.257 gives

$$\Delta((y,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{i_0\to k}) = 0$$

B. $\boldsymbol{(j<k)}$ then $((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0\to x_k})_k \underset{k\in\{1,\ldots,n\}\setminus\{i_o\}\Rightarrow k\neq i_0}{=} (y,x_1,\ldots,x_{j-1},x_{j+1},\ldots, x_n)_k = x_k$ and $((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0\to x_k})_{i_0} = x_k$, this together with $k\neq i_0$ and 10.257 gives

$$\Delta((y,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{i_0\to x_k}) = 0$$

Using (1) and (2) we find that $\Delta((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0\to x_k}) = \delta_{i,j} \cdot \Delta((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0\to x_k})$ and thus using 10.81 $\Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j = \sum_{k\in\{1,\ldots,n\}\setminus\{i_0\}} \delta_{j,k} \cdot (\alpha_k \cdot (\Delta((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0\to x_k}) \cdot x_j)) = \alpha_j \cdot \Delta((y,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{i_0\to x_j}) \cdot x_j$ giving that

$$\Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j = \alpha_j \cdot \Delta((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0\to x_j}) \cdot x_j \tag{10.82}$$

Now $(i_0 \rightsquigarrow j+1)\Delta((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0 \rightarrow x_j}) = \Delta((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0 \rightarrow x_j} \circ (i_0 \rightsquigarrow j+1)) \underset{j<i_0 \Rightarrow j+1 \leqslant i_0 \text{ and } 10.268}{=} \Delta(y, x_1, \ldots, x_{i_0-1}, x_{i_0+1}, \ldots, x_n)$ so using the skew symmetricity of Δ,10.96 and $j<i_0 \Rightarrow j \leqslant i_0+1$ we have that $(-1)^{i_0+1-j} \cdot \Delta((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0 \rightarrow x_j}) = \Delta(y, x_1, \ldots, x_{i_0-1}, x_{i_0+1}, \ldots x_n)$ and dividing both sides by $(-1)^{i_0+1-j}$ we get $\Delta((y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i_0 \rightarrow x_j}) = (-1)^{j-i_0-1}\Delta(y, x_1, \ldots, x_{i_0-1}, x_{i_0+1}, \ldots, x_n)$ and thus

$$\Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j = (-1)^{j-i_0-1} \cdot (\alpha_j \cdot (\Delta(y, x_1, \ldots, x_{i_0-1}, x_{i_0+1}, \ldots x_n) \cdot x_j)) \tag{10.83}$$

Using 10.76, 10.80 and 10.83 we have $B = \sum_{j \in \{1, \ldots, n\} \setminus \{i_0\}} (-1)^j \cdot \Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j = \sum_{j \in \{1, \ldots, i_0-1\}} (-1)^j \cdot \Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j + \sum_{j \in \{i_0+1, \ldots, n\}} (-1)^j \cdot \Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j = \sum_{j \in \{1, \ldots, i_0-1\}} (-1)^j((-1)^{j-i_0-1} \cdot (\alpha_j \cdot (\Delta(y, x_1, \ldots, x_{i_0-1}, x_{i_0+1}, \ldots x_n) \cdot x_j))) + \sum_{j \in \{i_0+1, \ldots, n\}} (-1)^j((-1)^{j-i_0-1} \cdot (\alpha_j \cdot (\Delta(y, x_1, \ldots, x_{i_0-1}, x_{i_0+1}, \ldots x_n) \cdot x_j))) = \sum_{j \in \{1, \ldots, n\} \setminus \{i_0\}} (-1)^j((-1)^{j-i_0-1} \cdot (\alpha_j \cdot (\Delta(y, x_1, \ldots, x_{i_0-1}, x_{i_0+1}, \ldots x_n) \cdot x_j))) = \Delta(y, x_1, \ldots, x_{i_0-1}, x_{i_0+1}, \ldots, x_n) \cdot \sum_{j \in \{1, \ldots n\} \setminus \{i_0\}} (-1)^{2j-i_0-1} \cdot (\alpha_j \cdot x_j) \underset{(-1)^{2j-i_0-1} = (-1)^{2j} \cdot (-1)^{-i_0} \cdot (-1)^{-1} = -(-1)^{i_0}}{=} -(-1)^{i_0} \cdot \Delta(y, x_1, \ldots, x_{i_0-1}, x_{i_0+1}, \ldots, x_n) \cdot \sum_{j \in \{1, \ldots, n\} \setminus \{i_0\}} \alpha_j \cdot x_j \underset{10.77}{=} -A$ so we have finally that $A + B = A + (-A) = 0$ and thus by 10.76 that $\hat{\Delta}(y, (x_1, \ldots, x_n)) = 0$ as should be

2. **($\{x_i\}_{i \in \{1, \ldots, n\}}$ are linear independent)** and thus as $\dim(X) = n$ we have by 10.157 and 10.199 that $\{x_i\}_{i \in \{1, \ldots, n\}}$ forms a basis of X. So if $y \in X$ there exists a unique $\{\alpha_i\}_{i \in \{1, \ldots, n\}}$ such that

$$y = \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot x_i \tag{10.84}$$

Then we have

$$\hat{\Delta}(y, (x_1, \ldots, x_n)) = \sum_{j \in \{1, \ldots, n\}} (-1)^{j-1} \cdot \Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j \tag{10.85}$$

and by multilinearity of Δ we have then that

$$\Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j = \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot \Delta(x_i, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j \tag{10.86}$$

Now for $\Delta(x_i, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$ we have if :

a. **($i < j$)** then $(x_i, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_1 = x_i = (x_i, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i+1}$ which as $1 \neq i+1$ would mean by 10.257 that $\Delta(x_i, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) = 0$.

b. $(\boldsymbol{j<i})$ then $(x_i, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_1 = x_i = (x_i, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_i$ which as $1 \leqslant j < i$ means by 10.257 that $\Delta(x_i, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) = 0$

So taking (a) and (b) in account we have $\Delta(x_i, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) = \delta_{i,j} \cdot \Delta(x_i, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$ and thus by 10.86 we have that $\Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j = \sum_{i\in\{1,\ldots,n\}} \delta_{i,j}(\alpha_i \cdot \Delta(x_i, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j) = \alpha_j \cdot (\Delta(x_j, x_1, \ldots., x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j)$ giving:

$$\Delta(y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j = \alpha_j \cdot (\Delta(x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \cdot x_j) \tag{10.87}$$

Now if $A = (x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \circ (1 \rightsquigarrow_n j)$ we have the following cases for $j \in \{1, \ldots, n\}$

a. $(\boldsymbol{j=1})$ then $(1 \rightsquigarrow_n j) = 1_{\{1,\ldots,n\}}$ so $A = (x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$ and $\forall k \in \{1, \ldots, n\}$ we have $(x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k = \begin{cases} x_j = x_1 \text{ if } k=1 \\ x_{k-1} \text{ if } k<j=1 \text{ which is impossible} \\ x_k \text{ if } k\in\{2,\ldots,n\} \end{cases} = x_k$ so $A = (x_1, \ldots, x_n)$.

b. $(\boldsymbol{j\in\{2,\ldots,n\}})$ then $1 < j$ and by definition of $(1 \sim_n j)$ for $k \in \{1, \ldots, n\}$ we have the following cases to consider for $A_k = ((x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \circ (1 \rightsquigarrow_n j))_k$:

i. $(\boldsymbol{k<1})$ this can not happen

ii. $(\boldsymbol{1 \leqslant k < j})$ then $A_k = (x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots x_n)_{k+1} \underset{2\leqslant k+1\leqslant j}{=} x_{(k+1)-1} = x_k$

iii. $(\boldsymbol{k=j})$ then $A_k = (x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_1 = x_j \underset{k=j}{=} x_k$

iv. $(\boldsymbol{j<k})$ then $A_k = (x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \underset{j<k}{=} x_k$

so in all cases we have $A_k = x_k$ so $A = (x_1, \ldots, x_n)$

As in (a) en (b) we have $A = (x_1, \ldots, x_n)$ so we have $(x_j, x_1, \ldots x_{j-1}, x_{j+1}, \ldots, x_n) \circ (1 \rightsquigarrow_n j) = (x_1, \ldots, x_n)$ and thus $(1 \rightsquigarrow_n j)\Delta(x_j, x_1, \ldots x_{j-1}, x_{j+1}, \ldots, x_n) = \Delta((x_j, x_1, \ldots x_{j-1}, x_{j+1}, \ldots, x_n) \circ (1 \rightsquigarrow_n j)) = \Delta(x_1, \ldots, x_n)$. Using the fact that Δ is skew symmetric and 10.96 we have then that $(-1)^{j-1} \cdot \Delta(x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) = \Delta(x_1, \ldots, x_n)$. Using the last equation, 10.85 and 10.87 we have then $\hat{\Delta}(y, (x_1, \ldots, x_n)) = \sum_{j\in\{1,\ldots,n\}} \alpha_j \cdot (\Delta(x_1, \ldots, x_n) \cdot x_j) = \Delta(x_1, \ldots, x_n) \cdot \sum_{j\in\{1,\ldots,n\}} \alpha_j \cdot x_j \underset{10.84}{=} \Delta(x_1, \ldots, x_n) \cdot y$.

So in all case (1) and (2) we have $\hat{\Delta}(y, (x_1, \ldots, x_n)) = \Delta(x_1, \ldots, x_n) \cdot y$ □

Definition 10.271. *Let $n \in \mathbb{N}$, X be a n-dimensional vector space over a field F, Δ a non trivial determinant function $\Delta: X^n \to F$ and $L: X \to X$ a linear mapping ($L \in \operatorname{Hom}(X, X)$) (we call this sometimes a linear transformation). Then Δ_L: $X^n \to F$ is defined by $(x_1, \ldots, x_n) \to \Delta_L(x_1, \ldots, x_n) = \Delta(L(x_1), \ldots, L(x_n))$*

Theorem 10.272. *Let $n \in \mathbb{N}$, X be a n-dimenional vector spaces over a field F of characterization zero, Δ a non trivial determinant function $\Delta: X^n \to F$ in X and $L \in \operatorname{Hom}(X, X)$ a linear transformation then Δ_L is a determinant function.*

Proof.

1. **(Multilinearity)**

$$\begin{aligned}\Delta_L(x_1, \ldots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_n) &= \Delta(L(x_1), \ldots, L(x_{i-1}), L(\alpha \cdot x + \beta \cdot y), L(x_{i+1}), \ldots, L(x_n)) \\ &\underset{L \text{ is linear}}{=} \Delta(L(x_1), \ldots, L(x_{i-1}), \alpha \cdot L(x) + \beta \cdot L(x), L(x_{i+1}), \ldots, L(x_n)) \\ &\underset{\Delta \text{ is multilinear}}{=} \alpha \cdot \Delta(L(x_1), \ldots, L(x_{i-1}), L(x), L(x_{i+1}), \ldots, L(x_n)) + \beta \cdot \Delta(L(x_1), \ldots, L(x_{i-1}), L(y), L(x_{i+1}), \ldots, L(x_n)) \\ &= \alpha \cdot \Delta_L(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) + \beta \cdot \Delta_L(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)\end{aligned}$$

2. **(Skew Symmetry)** Assume that we have $(x_1, \ldots, x_n) \in X^n$ such that $\exists i, j \in \{1, \ldots, n\}$ with $i \neq j$ and $x_i = x_j$ then $L(x_i) = L(x_j)$ then $\Delta_L(x_1, \ldots, x_n) = \Delta(L(x_1), \ldots L(x_n)) \underset{\Delta \text{ is skew symmetric and } 10.257}{=} 0$ so by 10.257 we have that Δ_L is skew symmetric. □

Theorem 10.273. *Let $n \in \mathbb{N}$, X be a n-dimenional vector spaces over a field F of characterization zero, Δ a non trivial determinant function $\Delta: X^n \rightarrow F$ in X and $L \in \operatorname{Hom}(X, X)$ a linear transformation then there exists a unique scalar $\alpha \in F$ such that $\Delta_L = \alpha \cdot \Delta$ (unique meaning that α is only determined by L not by Δ) or if $\Delta': X^n \rightarrow F$ is another non trivial determinant function then $\Delta'_L = \alpha \cdot \Delta'$.*

Proof.

1. **(existence)** Using 10.265 and the fact that by the above theorem Δ_L is determinant function we have the existence of a $\alpha \in F$ such that $\Delta_L = \alpha \cdot \Delta$.
2. **(uniqueness)** If Δ' is another non trivial determinant function then by 10.265 there exists a α' such that $\Delta' = \lambda \cdot \Delta$. So if $(x_1, \ldots, x_n) \in X^n$ then $\Delta'_L(x_1, \ldots, x_n) = \Delta'(L(x_1), \ldots, L(x_n)) = \lambda \cdot \Delta(L(x_1), \ldots, L(x_n)) = \lambda \cdot \Delta_L(x_1, \ldots, x_n) \Rightarrow \Delta'_L = \lambda \cdot \Delta_L = \lambda \cdot (\alpha \cdot \Delta) = \alpha \cdot (\lambda \cdot \Delta) = \alpha \cdot \Delta' \Rightarrow \Delta'_L = \alpha \cdot \Delta'$ if now also $\Delta'_L = \alpha' \cdot \Delta'$ then by 10.264 we must have $\alpha = \alpha'$. □

Now by 10.263 there always exists a non trivial determinant function so that the following definition make sense (together with the above theorem)

Definition 10.274. *Let $n \in \mathbb{N}, X$ be a n-dimenional vector spaces over a field F of characterization zero, $L: X \rightarrow X$ be a linear transformation [$L \in \operatorname{Hom}(X, X)$]. Then $\det(L)$ is defined so that $\Delta_L = \det(L) \cdot \Delta$ where Δ is any non trivial determinant function (by 10.263 such a determinant function exists and by the previous theorem $\det(X)$ is independent of the determinant function). So $\det: \operatorname{Hom}(X, X) \rightarrow F$ defined by $L \rightarrow \det(L)$ is a well defined function.*

Example 10.275. Let X be a n-dimenional vector space over a field F of characterization zero then if $\lambda \in F$ we have that $\lambda \cdot 1_X: X \rightarrow X$ is trivially a linear transformations. Now if Δ is a trivial determinant function we have that $\Delta_{\lambda \cdot 1_X}(x_1, \ldots, x_n) = \Delta(\lambda \cdot 1_X(x_1), \ldots, \lambda \cdot 1_X(x_n)) = \Delta(\lambda \cdot x_1, \ldots, \lambda \cdot x_n) = \lambda^n \cdot \Delta(x_1, \ldots, x_n)$ so we have that $\mathbf{det(\boldsymbol{\lambda} \cdot 1_X) = \boldsymbol{\lambda}^n}$. Or if we take $\lambda = 1$ (the unit in F) that $\mathbf{det(1_X) = 1}$.

Example 10.276. If X is a 1 dimensional space then if $\{e_1\}$ is the basis of X then if $x_1 = \alpha \cdot e_1 \in X$ we have for $L \in \operatorname{Hom}(X, X)$ that $L(x_1) = \alpha \cdot L(e_1)$, as $L(e_1) \in X$ there exists a β such that $L(e_1) = \beta \cdot e_1$. So we have $\Delta_L(x_1) = \Delta(L(x_1)) = \Delta(\alpha \cdot L(e_1)) = \Delta(\alpha \cdot \beta \cdot e_1) = \beta \cdot \Delta(\alpha \cdot e_1) = \beta \cdot \Delta(x)$ or we have that $\det(L) = \beta$ (where $L(e_1) = \beta \cdot e_1$).

Example 10.277. If X is a n-dimensional space with basis $\{e_i\}_{i \in \{1, \ldots, n\}}$ over a field F of characterization zero, $\alpha \in F$, $i \in \{1, \ldots, n\}$ and $L: X \rightarrow X$ defined by $L(e_k) = \begin{cases} e_k \text{ if } k \neq i \\ \alpha \cdot e_k \text{ if } k = i \end{cases}$ then $\det(L) = \alpha$

Proof. $\Delta_L(e_1, \ldots, e_n) = \Delta(L(e_1), \ldots L(e_{i-1}), L(e_i), L(e_{i+1}), \ldots, L(e_n)) = \Delta(e_1, \ldots e_{i-1}, \alpha \cdot e_i, e_{i+1}, \ldots, e_n) = \alpha \cdot \Delta(e_1, \ldots, e_n)$ proving that $\det(L) = \alpha$ □

Example 10.278. If X is a n-dimensional space with basis $\{e_i\}_{i \in \{1, \ldots, n\}}$ over a field F of characterization zero, $i, j \in \{1, \ldots, n\}$ with $i \neq j$ and $L: X \rightarrow X$ defined by $L(e_k) = \begin{cases} e_k \text{ if } k \neq i, j \\ e_i \text{ if } k = j \\ e_j \text{ if } k = i \end{cases} \Rightarrow L(e_k) = e_{(i \leftrightarrow_n j)(k)}$ then $\det(L) = -1$

Proof. $\Delta_L(e_1, \ldots, e_n) = \Delta(L(e_1), \ldots, L(e_n)) = \Delta(e_{(i \leftrightarrow_n j)}, \ldots, e_{(i \leftrightarrow_n j)(n)}) = \operatorname{sign}((i \leftrightarrow_n j)) \cdot \Delta(e_1, \ldots, e_n) = (-1) \cdot \Delta(e_1, \ldots, e_n)$ proving that $\det(L) = -1$ □

Example 10.279. If X is a n-dimensional space with basis $\{e_i\}_{i \in \{1, \ldots, n\}}$ over a field of characterization zero, let $i, j \in \{1, \ldots, n\} \vdash i \neq j$ and $\alpha \neq 0$ define then L by $L(e_k) = \begin{cases} e_k \text{ if } k \neq i \\ e_i + \alpha \cdot e_j \text{ if } k = i \end{cases}$ then $\det(L) = 1$

Proof.

$$\begin{aligned}
\Delta_L(e_1, \ldots, e_n) &= \Delta(L(e_1), \ldots, L(e_{i-1}), L(e_i), L(e_{i+1}), \ldots, L(e_n)) \\
&= \Delta(e_1, \ldots, e_{i-1}, e_i + \alpha \cdot e_j, e_{i+1}, \ldots, e_n) \\
&= \Delta(e_1, \ldots, e_{i-1}, e_i, e_{i+1}, \ldots, e_n) + \alpha \cdot \Delta(e_1, \ldots, e_{i-1}, e_j, e_{i+1}, \ldots e_n) \\
&\underset{10.257}{=} \Delta(e_1, \ldots, e_n) + \alpha \cdot 0 \\
&= \Delta(e_1, \ldots, e_n)
\end{aligned}$$

proving that $\det(L) = 1$ □

Theorem 10.280. *Let $n \in \mathbb{N}$, X be a n-dimenional vector spaces over a field F of characterization zero then we have*

1. $\det(1_X) = 1$ *(see for the proof the previous example)*

2. *A linear transformation* $L: X \rightarrow X$ *is regular (meaning injective and thus a isomorphism by 10.212) if and only if* $\det(L) \neq 0$

3. *If* $L_1, L_2: X \rightarrow X$ *are linear transformations then* $\det(L_1 \circ L_2) = \det(L_1) \cdot \det(L_2)$. *From this if follows that if* $L: X \rightarrow X$ *is a bijective linear transformation (a linear isomorphism) then* $L \circ L^{-1} = 1_X \Rightarrow \det(L) \cdot \det(L^{-1}) = \det(1_X) = 1 \Rightarrow \det(L^{-1}) = \det(L)^{-1}$

Proof.

1. This was proved in the previous example.

2. First we prove that regularity implies that $\det(L) \neq 0$. If $\{e_i\}_{i \in \{1,\ldots,n\}}$ is a basis of X then as L is a regular linear transformation we can use 10.212 and 10.207 to prove that $\{L(e_i)\}_{i \in \{1,\ldots,n\}}$ is a basis of X. Then using the determinant function defined in 10.263 we have that

$$\begin{aligned} 1 &= \Delta(L(e_1), \ldots, L(e_n)) \\ &= \Delta_L(e_1, \ldots, e_n) \\ &= \det(L) \cdot \Delta(e_1, \ldots, e_n) \end{aligned}$$

If now $\det(L) = 0$ then we have $1 = 0$ which is a contradiction as F is field of characterization zero, so

$$\det(L) \neq 0$$

Next we prove that if $\det(L) \neq 0$ then L is regular. Using 10.263 find a Δ' such that

$$\Delta'(e_1, \ldots, e_n) = 1$$

then

$$\begin{aligned} \Delta'(L(e_1), \ldots, L(e_n)) &= \det(L) \cdot \Delta'(e_1, \ldots, e_n) \\ &= \det(L) \\ &\neq 0 \end{aligned}$$

So we must have that $\{L(e_i)\}_{i \in \{1,\ldots,n\}}$ is linear independent [if not we have by 10.257 that $\Delta'(L(e_1), \ldots, L(e_n)) = 0$]. Using 10.212 we have then that L is regular.

3. If $(x_1, \ldots x_n) \in X^n$ and Δ any non trivial determinant function (which must exists by 10.263) then we have

$$\begin{aligned} \Delta_{L_1 \circ L_2}(x_1, \ldots, x_n) &= \Delta((L_1 \circ L_2)(x_1), \ldots, (L_1 \circ L_2)(x_n)) \\ &= \Delta(L_1(L_2(x_1)), \ldots, L_1(L_2(x_n))) \\ &= \Delta_{L_1}(\Delta(L_2(x_1)), \ldots, L_2(x_n)) \\ &= \det(L_1) \cdot \Delta(L_2(x_1), \ldots, L_2(x_n)) \\ &= \det(L_1) \cdot \Delta_{L_2}(x_1, \ldots, x_n) \\ &= (\det(L_1) \cdot \det(L_2)) \cdot \Delta(x_1, \ldots, x) \end{aligned}$$

and thus by definition of $\det(L_1 \circ L_2)$ we have that $\det(L_1 \circ L_2) = \det(L_1) \cdot \det(L_2)$. $\square$

Definition 10.281. *Let $n \in \mathbb{N}$, X be a n-dimensional vector space over a field F of characterization zero, Δ a non trivial determinant function in X, $L \in \operatorname{Hom}(X, X)$ then we define $\breve{\Delta}_L: X^n \to \operatorname{Hom}(X, X)$ by $(x_1, \ldots, x_n) \to \breve{\Delta}_L(x_1, \ldots, x_n)$ where $\breve{\Delta}_L(x_1, \ldots, x_n): X \to X$ is defined by $x \to \breve{\Delta}_L(x_1, \ldots, x_n)(x) = \sum_{i \in \{1, \ldots, n\}} (-1)^{i-1} \cdot \Delta(x, L(x_1), \ldots, L(x_{i-1}), L(x_{i+1}), \ldots, L(x_n)) \cdot x_i$*

Proof. Of course we must prove that $\breve{\Delta}_L(x_1, \ldots, x_n)$ is a linear mapping. So let $\alpha, \beta \in F$, $x, y \in X$ and $(x_1, \ldots, x_n) \in X^n$ then we have

$$\begin{aligned}
\breve{\Delta}_L(x_1, \ldots, x_n)(\alpha \cdot x + \beta \cdot y) &= \sum_{i \in \{1, \ldots, n\}} (-1)^{i-1} \cdot \Delta(\alpha \cdot x + \beta \cdot y, L(x_1), \ldots, L(x_{i-1}), L(x_{i+1}), \ldots, L(x_n)) \cdot x_i \\
&\underset{L \text{ is multilinear}}{=} \sum_{i \in \{1, \ldots, n\}} (-1)^{i-1} \cdot (\alpha \cdot \Delta(x, L(x_1), \ldots, L(x_{i-1}), L(x_{i+1}), \ldots, L(x_n)) + \beta \cdot \Delta(y, L(x_1), \ldots, L(x_{i-1}), L(x_{i+1}), \ldots, L(x_n))) \cdot x_i \\
&= \alpha \cdot \sum_{i \in \{1, \ldots, n\}} (-1)^{i-1} \cdot \Delta(x, L(x_1), \ldots, L(x_{i-1}), L(x_{i+1}), \ldots, L(x_n)) \cdot x_i + \beta \cdot \sum_{i \in \{1, \ldots, n\}} (-1)^{i-1} \cdot \Delta(y, L(x_1), \ldots, L(x_n)) \cdot x_i \\
&= \alpha \cdot \breve{\Delta}_L(x_1, \ldots, x_n)(x) + \beta \cdot \breve{\Delta}_L(x_1, \ldots, x_n)(y)
\end{aligned}$$

$\square$

Example 10.282.

1. If $n = 1$ then $\breve{\Delta}_L(x_1)(x) = (-1)^{1-1} \cdot \Delta(x) \cdot x_1 = \Delta(x) \cdot x_1$

2. If $n = 2$ then $\breve{\Delta}_L(x_1, x_2) = (-1)^{1-1} \cdot \Delta(x, L(x_2)) \cdot x_1 + (-1)^{2-1} \cdot \Delta(x, L(x_1)) \cdot x_2 = \Delta(x, L(x_2)) \cdot x_1 - \Delta(x, L(x_1)) * x_2$

Lemma 10.283. *Let $n \in \mathbb{N}$, X as set and $x = (x_1, \ldots, x_n) \in X^n$, $i, j \in \{1, \ldots, n\}$ with $i \neq j$ and $x_i = x_j$ and $t \in X$ then we have the following*

1. ***(j < i)*** *then $(t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \circ (i \rightsquigarrow_n j + 1) = (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$*

2. ***(i < j)*** *then $(t, x_1, \ldots x_{j-1}, x_{j+1}, \ldots, x_n) \circ (i + 1 \rightsquigarrow_n j) = (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$*

Proof.

1. $(\boldsymbol{j<i})$ then $j+1\leqslant i$ let then $a=(t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)\circ(i\rightsquigarrow_n j+1)$ then we have the following to consider:

   a. $(\boldsymbol{j+1=i})$ then $a=(t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)$ and we have for $k\in\{1,\ldots,n\}$ that

      i. $(\boldsymbol{k=1})$ then $a_k=t=(t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k$

      ii. $(\boldsymbol{2\leqslant k\leqslant j})$ then $a_k=x_{k-1}\underset{k\leqslant j<i}{=}(t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k$

      iii. $(\boldsymbol{k=j+1})$ then $a_k=x_k=x_i=x_j=x_{k-1}\underset{j+1=i}{=}(t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k$

      iv. $(\boldsymbol{j+1<k})$ then $a_k=x_k\underset{i=j+1<k}{=}(t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k$

   this proves that $(t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)\circ(i\sim_n j+1)=(t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)$.

   b. $(\boldsymbol{j+1<i})$ then we have by the definition of $(i\rightsquigarrow_n j+1)$ the following cases to consider for $k\in\{1,\ldots,n\}$ in $a=(t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)\circ(i\rightsquigarrow_n j+1)$:

      i. $(\boldsymbol{k=1})$ then as $k=1<j+1$ we have

$$\begin{aligned}a_k &= (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k\\ &\underset{k=1}{=} t\\ &\underset{k=1}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k\end{aligned}$$

      ii. $(\boldsymbol{k\neq 1\wedge k<j+1})$ then

$$\begin{aligned}a_k &= (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k\\ &\underset{k<j+1<i}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k\end{aligned}$$

      iii. $(\boldsymbol{k\neq 1\wedge k=j+1})$ then

$$\begin{aligned}a_k &= (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_i\\ &\underset{j+1<i}{=} x_i\\ &\underset{x_i=x_j}{=} x_j\\ &\underset{j+1<i}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{j+1}\\ &\underset{k=j+1}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k\end{aligned}$$

      iv. $(\boldsymbol{k\neq 1\wedge j+1<k\leqslant i})$ then

$$\begin{aligned}a_k &= (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{k-1}\\ &\underset{k\leqslant i}{=} (t,x+1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k\end{aligned}$$

v. $(\boldsymbol{k \neq 1 \wedge i < k})$ then

$$\begin{array}{rcl} a_k & = & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \\ & j < j+1 < i < k & x_k \\ & \underset{i<k}{=} & (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \end{array}$$

this proves that $(t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \circ (i \sim_n j+1) = (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$.

2. $(\boldsymbol{i < j})$ then $i+1 \leqslant j$ and we have by the definition of $(i+1 \rightsquigarrow_n j)$ the following cases to consider for $a = (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \circ (i+1 \rightsquigarrow_n j)$

   a. $(\boldsymbol{i+1 = j})$ then we have that $a = (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$ and we have for $k \in \{1, \ldots, n\}$ that :

      i. $(\boldsymbol{k = 1})$ then $a_k = t = (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k$

      ii. $(\boldsymbol{k \leqslant i})$ then as $k \leqslant i < j$ we have $a_k = x_{k-1} \underset{k \leqslant i}{=} (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k$

      iii. $(\boldsymbol{k = i+1 = j})$ then as $k = j \leqslant j$ we have $a_k = x_{k-1} \underset{k=i+1 \Rightarrow k-1=i}{=} x_i \underset{x_i = x_j}{=} x_j = x_k \underset{i<i+1=k}{=} (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k$

      iv. $(\boldsymbol{j < k})$ then we have $a_k = x_k \underset{i<j<k}{=} (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k$

   this proves that $(t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \circ (i+1 \rightsquigarrow_n j) = (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$

   b. $(\boldsymbol{i+1 < j})$ then we have by the definition of $(i+1 \rightsquigarrow_n j)$ the following cases to consider for $k \in \{1, \ldots, n\}$ in $a = (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \circ (i \rightsquigarrow_n j+1)$:

      i. $(\boldsymbol{k = 1})$ then

$$\begin{array}{rcl} a_k & \underset{1<i+1}{=} & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \\ & \underset{k=1}{=} & t \\ & \underset{k=1}{=} & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \end{array}$$

      ii. $(\boldsymbol{k \neq 1 \wedge k < i+1})$ then

$$\begin{array}{rcl} a_k & = & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \\ & \underset{k<i+1<j}{=} & x_{k-1} \\ & \underset{k<i+1 \Rightarrow k \leqslant i}{=} & (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \end{array}$$

      iii. $(\boldsymbol{k \neq 1 \wedge i+1 \leqslant k < j})$ then

$$\begin{array}{rcl} a_k & = & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{k+1} \\ & \underset{k<j \Rightarrow k+1 \leqslant j}{=} & x_k \\ & \underset{i<i+1 \leqslant k}{=} & (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \end{array}$$

iv. $(\boldsymbol{k \neq 1 \wedge k = j})$ then

$$\begin{array}{ccl} a_k & = & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i+1} \\ & \underset{i+1<j}{=} & x_i \\ & \underset{x_i = x_j}{=} & x_j \\ & \underset{k=j}{=} & x_k \\ & \underset{i+1<j=k \Rightarrow i<k}{=} & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \end{array}$$

v. $(\boldsymbol{k \neq 1 \wedge j < k})$ then

$$\begin{array}{ccl} a_k & = & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \\ & \underset{j<k}{=} & x_k \\ & i < i+1 < j < k & (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \end{array}$$

this proves that $(t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \circ (i+1 \rightsquigarrow_n j) = (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$

□

Theorem 10.284. *Let $n \in \mathbb{N}$, X be a n-dimensional vector space over a field F of characterization zero, Δ a non trivial determinant function in X, $L \in \operatorname{Hom}(X, X)$ then $\breve{\Delta}_L: X^n \rightarrow \operatorname{Hom}(X, X)$ is n-linear and skew symmetric.*

Proof. Let $i \in \{1, \ldots, n\}$, $\alpha, \beta \in F$ and $x, y \in X$, $(x_1, \ldots, x_n) \in X^n$ such that $x_i = \alpha \cdot x + \beta \cdot y$ then we have for $t \in X$, define then $\forall j \in \{1, \ldots, n\}$

$$A_j(x_1, \ldots x_n) = \Delta(t, L(x_1), \ldots, L(x_{j-1}), L(x_{j+1}), \ldots, L(x_n)) \cdot x_j \tag{10.88}$$

then we have by definition that

$$\breve{\Delta}_L(x_1, \ldots, x_n)(t) = \sum_{j \in \{1, \ldots, n\}} (-1)^{j-1} \cdot A_j(x_1, \ldots, x_n) \tag{10.89}$$

We have now the following cases for j to consider

1. $(\boldsymbol{j = i})$ then

$$\begin{array}{rcl} A_j(x_1, \ldots, x_n) & = & \Delta(t, L(x_1), \ldots, L(x_{i-1}), L(x_{i+1}), \ldots, L(x_n)) \cdot (\alpha \cdot x + \beta \cdot y) \\ & = & \alpha \cdot \Delta(t, L(x_1), \ldots, L(x_{i-1}), L(x_{i+1}), \ldots, L(x_n)) \cdot x + \beta \cdot \Delta(t, \\ & & L(x_1), \ldots, L(x_{i-1}), L(x_{i+1}), \ldots, L(x_n)) \cdot x \\ & = & \alpha \cdot A_j((x_1, \ldots, x_n)_{i \rightarrow x}) + \beta \cdot A_j((x_1, \ldots, x_n)_{i \rightarrow y}) \end{array}$$

2. $(\boldsymbol{j < i})$ then $(t, L(x_1), \ldots, L(x_{j-1}), L(x_{j+1}), \ldots, L(x_n))_i = L(x_i) = L(\alpha \cdot x + \beta \cdot y) = \alpha \cdot L(x) + \beta \cdot L(y)$ and thus by multi linearity of Δ we have

$$\begin{array}{rcl} A_j(x_1, \ldots, x_n) & = & \Delta(t, L(x_1), \ldots, L(x_{j-1}), L(x_j), \ldots, L(x_n)) \\ & = & \alpha \cdot \Delta((t, L(x_1), \ldots, L(x_{j-1}), L(x_j), \ldots L(x_n))_{i \rightarrow L(x)}) + \beta \cdot \\ & & \Delta((t, L(x_1), \ldots, L(x_{j-1}), L(x_j), \ldots L(x_n))_{i \rightarrow L(y)}) \\ & = & \alpha \cdot A_j((x_1, \ldots, x_n)_{i \rightarrow x}) + \beta \cdot A_j((x_1, \ldots, x_n)_{i \rightarrow y}) \end{array}$$

3. $(\boldsymbol{i<j})$ then $(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))_{i+1}=L(x_i)=L(\alpha\cdot x+\beta\cdot y)=\alpha\cdot L(x)+\beta\cdot L(y)$ and thus by multi linearity of Δ we have

$$\begin{aligned}A_j(x_1,\ldots,x_n) &= \Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\\ &= \alpha\cdot\Delta((t,L(x_1),\ldots,L(x_{j-1}),L(x_j),\ldots L(x_n))_{i\to L(x)})+\beta\cdot\\ &\quad \Delta((t,L(x_1),\ldots,L(x_{j-1}),L(x_j),\ldots L(x_n))_{i\to L(y)})\\ &= \alpha\cdot A_j((x_1,\ldots,x_n)_{i\to x})+\beta\cdot A_j((x_1,\ldots,x_n)_{i\to y})\end{aligned}$$

so taken (1),(2) and (3) in account we have by 10.89 that

$$\begin{aligned}\breve{\Delta}_L(x_1,\ldots,x_n)(t) &= \alpha\cdot\sum_{j\in\{1,\ldots,n\}}(-1)^{j-1}\cdot A_j((x_1,\ldots,x_n)_{i\to x})+\beta\cdot\\ &\quad\sum_{j\in\{1,\ldots,n\}}(-1)^{j-1}\cdot A_j((x_1,\ldots,x_n)_{i\to y})\\ &= \alpha\cdot\breve{\Delta}_L((x_1,\ldots,x_n)_{i\to x})+\beta\cdot\breve{\Delta}_L((x_1,\ldots,x_n)_{i\to y})\end{aligned}$$

proving that $\breve{\Delta}_L$ is indeed multi linear.

To prove that $\breve{\Delta}_L$ is skew symmetric let $x=(x_1,\ldots,x_n)\in X^n$, $t\in X$ such that there exists a $i_0,j_0\in\{1,\ldots,n\}$ with $i_0\neq j_0$ and $x_{i_0}=x_{j_0}$ then we have the following cases to consider for $\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))$

1. $(\boldsymbol{i_0<j_0})$ look then at the following cases for $j\in\{1,\ldots,n\}$

   a. $(\boldsymbol{j\neq i_0,j_0})$ then we have the following sub cases to consider

      i. $(\boldsymbol{j<i_0<j_0})$ then $(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1})\ldots,L(x_n))_{i_0}=L(x_{i_0})\underset{x_{i_0}=x_{j_0}}{=}L(x_{j_0})=(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))_{j_o}$ so that using the skew symmetry of Δ we have $\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))=0$

      ii. $(\boldsymbol{i_0<j<j_0})$ then $(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1})\ldots,L(x_n))_{i_0+1}=L(x_{i_0})\underset{x_{i_0}=x_{j_0}}{=}L(x_{j_0})=(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))_{j_o}$ so that using the skew symmetry of Δ we have $\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))=0$

      iii. $(\boldsymbol{i_0<j_0<j})$ then $(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1})\ldots,L(x_n))_{i_0+1}=L(x_{i_0})\underset{x_{i_0}=x_{j_0}}{=}L(x_{j_0})=(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))_{j_o+1}$ so that using the skew symmetry of Δ we have $\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))=0$

   so we have

$$\text{if } j\neq i_0,j_0\Rightarrow\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))=0 \qquad (10.90)$$

   b. $(\boldsymbol{j=i_0})$ for $j=i_0<j_0$ so using the above lemma together with $L(x_{i_0})=L(x_{j_0})$ we have that $(t,L(x_1),\ldots,L(x_{i_0-1}),L(x_{i_0+1}),\ldots,L(x_n))\circ(j_0\rightsquigarrow_n i_0+1)=(t,L(x_1),\ldots,L(x_{j_0-1}),L(x_{j_0+1}),\ldots,L(x_n))$, using the skew symmetry of Δ and 10.96 we have that $(-1)^{j_0-i_0-1}\Delta(t,L(x_1),\ldots,L(x_{i_0-1}),L(x_{i_0+1}),\ldots,L(x_n))=(j_0\rightsquigarrow_n i_o+1)\Delta(t,L(x_1),\ldots L(x_{i_0-1}),$

$L(x_{i_0+1}),\ldots,L(x_n))=\Delta((t,L(x_1),\ldots,L(x_{i_0-1}),L(x_{i_0+1}),\ldots,L(x_n))\circ(j_0\rightsquigarrow_n i_0+1))=\Delta(t,L(x_1),\ldots,L(x_{j_0-1}),L(x_{j_0+1}),\ldots,L(x_n))$ giving thus

$$(-1)^{j_0-i_0-1}\Delta(t,\ L(x_1),\ \ldots,\ L(x_{i_0-1}),\ L(x_{i_0+1}),\ \ldots,\ L(x_n))=\Delta(t,\ L(x_1),\ldots,L(x_{j_0-1}),L(x_{j_0+1}),\ldots,L(x_n)) \tag{10.91}$$

Now

$$\begin{aligned}
\breve{\Delta}_L(x_1,\ldots,x_n)(t) &= \sum_{j\in\{1,\ldots,n\}}(-1)^{j-1}\cdot\Delta(t,\ L(x_1),\ \ldots,\ L(x_{j-1}),\ L(x_{j+1}),\ldots,L(x_n))\cdot x_j\\
&= \sum_{j\in\{1,\ldots,n\}\setminus\{i_0,j_0\}}(-1)^{j-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\cdot x_j+\sum_{j\in\{i_0\}}(-1)^{j-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\cdot x_j+\\
&\quad \sum_{j\in\{j_0\}}(-1)^{j-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\cdot x_j\\
&\underset{10.90}{=} \sum_{j\in\{1,\ldots,n\}\setminus\{i_0,j_0\}}(-1)^{j-1}\cdot 0\cdot x_j+(-1)^{i_0-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{i_0-1}),L(x_{i_0+1}),\ldots,L(x_n))\cdot x_{i_0}+\\
&\quad (-1)^{j_0-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{j_0-1}),L(x_{j_0+1}),\ldots,L(x_n))\cdot x_{j_0}\\
&\underset{10.91}{=} (-1)^{i_0-1}\cdot(-1)^{j_0-i_0-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{j_0-1}),\ldots,L(x_n))\cdot x_{i_0}+(-1)^{j_0-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{j_0-1}),L(x_{j_0+1}),\ldots,L(x_n))\cdot x_{j_0}\\
&\underset{x_{i_0}=x_{j_0}}{=} (-1)^{j_0-1}\cdot(-1)\Delta(t,\ L(x_1),\ \ldots,\ L(x_{j_0-1}),\ \ldots,L(x_n))\cdot x_{j_0}+(-1)^{j_0-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{j_0-1}),L(x_{j_0+1}),\ldots,L(x_n))\cdot x_{j_0}\\
&= 0
\end{aligned}$$

2. $(\boldsymbol{j_0<i_0})$ take then $i_0'=j_0$ and $j_0'=i_0$ then $i_0'<j_0'$ and by using (1) we have that $\breve{\Delta}_L(x_1,\ldots,x_n)=0$

So in (1) and (2) we have proved that $\breve{\Delta}_L(x_1,\ldots,x_n)=0$ which means by 10.257 that $\breve{\Delta}_L$ is skew symmetric.

□

Now using the fact that by the above theorem $\breve{\Delta}_L: X^n\to\mathrm{Hom}(X,X)$ is n-linear and skew symmetric together with 10.264 there exists a unique $y\in\mathrm{Hom}(X,X)$ such that $\breve{\Delta}_L(x_1,\ldots,x_n)=\Delta(x_1,\ldots,x_n)\cdot y$. We call this unique y the adjoint of L and note it by adjoint(L). This leads directly to the following definition.

Definition 10.285. *Let $n \in \mathbb{N}$, X be a n-dimensional vector spaces over a field F with characteristics zero, $L \in \operatorname{Hom}(X, X)$ a linear transformation. Then there exists a unique (it does only dependent on L)* $\operatorname{adjoint}(L) \in \operatorname{Hom}(X, X)$ *such that for every non trivial determinant function Δ we have $\breve{\Delta}_L(x_1, \ldots, x_n) = \Delta(x_1, \ldots, x_n) \cdot \operatorname{adjoint}(L)$*

Proof. To prove uniqueness let Δ' be another non trivial determinant function and $\operatorname{adjoint}'(L)$ such that $\breve{\Delta'}_L(x_1, \ldots, x_n) = \Delta'(x_1, \ldots, x_n) \cdot \operatorname{adjoint}'(L)$. Now using 10.265 there exists a $\lambda \in F$ such that $\Delta' = \lambda \cdot \Delta$. Then we have

$$
\begin{aligned}
\Delta'(x_1, \ldots, x_n) \cdot \operatorname{adjoint}'(L)(t) &= \breve{\Delta'}_L(x_1, \ldots, x_n)(t) \\
&= \sum_{j \in \{1, \ldots, n\}} (-1)^{j-1} \cdot \Delta'(t, L(x_1), \ldots, L(x_{j-1}) \cdot \\
&\quad L(x_{j+1}), \ldots, L(x_n)) \cdot x_j \\
&= \sum_{j \in \{1, \ldots, n\}} (-1)^{j-1} \cdot \lambda \cdot \Delta(t, L(x_1), \ldots, L(x_{j-1}) \cdot \\
&\quad L(x_{j+1}), \ldots, L(x_n)) \cdot x_j \\
&= \lambda \cdot \sum_{j \in \{1, \ldots, n\}} (-1)^{j-1} \cdot \Delta(t, L(x_1), \ldots, L(x_{j-1}) \cdot \\
&\quad L(x_{j+1}), \ldots, L(x_n)) \cdot x_j \\
&= \lambda \cdot \breve{\Delta}_L(x_1, \ldots, x_n)(t) \\
&= \lambda \cdot (\Delta(x_1, \ldots, x_n) \cdot \operatorname{adjoint}(L)(t)) \\
&= \Delta'(x_1, \ldots, x_1) \cdot \operatorname{adjoint}(L)(t)
\end{aligned}
$$

As Δ' is non trivial there exists a $(x_1, \ldots, x_n)$ such that $\Delta'(x_1, \ldots, x_n) \neq 0$ then by dividing by $\Delta'(x_1, \ldots, x_n)$ we get $\operatorname{adjoint}'(L) = \operatorname{adjoint}(L)$. □

Example 10.286. If $n = 1$ then if $\{e_1\}$ is a basis of X then if $x \in X$ we have $x = \alpha \cdot e_1$ and $L(x) = \alpha \cdot L(e_1)$, from 10.282 we have that $\breve{\Delta}_L(x_1)(x) = \Delta(x) \cdot x_1$ and if $x_1 = \beta \cdot e_1$ we have $\breve{\Delta}_L(x_1)(x) = \alpha \cdot \beta \cdot \Delta(e_1) \cdot e_1 = \Delta(\beta \cdot e_1) \cdot (\alpha \cdot e_1) = \Delta(x_1) \cdot x$ so that $\operatorname{adjoint}(L)(x) = x$ or we have that $\operatorname{adjoint}(L) = i_X$. Note also that if $L(e_1) = \gamma \cdot e_1$ we have by 10.276 that $\det(L) = \gamma$ so that we have

1. $\operatorname{adjoint}(L(x)) = L(x) = L(\alpha \cdot e_1) = \alpha \cdot L(e_1) = \alpha \cdot (\gamma \cdot e_1) = \gamma \cdot (\alpha \cdot e_1) = \gamma \cdot x = \det(L) \cdot x = \det(L) \cdot 1_X(x)$ or $\operatorname{adjoint}(L) \circ L = \det(L) \cdot 1_X$

2. $L(\operatorname{adjoint}(x)) = L(x) \underset{\text{see}(1)}{=} \det(L) \cdot 1_X(x)$ or $L \circ \operatorname{adjoint}(L) = \det(L) \cdot 1_X$

Actually we can generalize this result to any dimension as the following theorem shows

Theorem 10.287. *Let $n \in \mathbb{N}$, X be a n-dimensional vector space over a field of characterization zero, $L \in \operatorname{Hom}(X, X)$ then we have*

1. $\operatorname{adjoint}(L) \circ L = \det(L) \cdot 1_X$

2. $L \circ \operatorname{adjoint}(L) = \det(L) \cdot 1_X$

Proof.

1. *Let* $(e_1, \ldots, e_n)$ *a basis for* X *and let* Δ *the non zero determinant function defined in* 10.263 *such that* $\Delta(e_1, \ldots, e_n) = 1$ *then if* $x \in X$ *we have*

$$\begin{array}{rcl}(\mathrm{adjoint}(L) \circ L)(x) & = & \mathrm{adjoint}(L)(L(x)) \\ & = & \Delta(e_1, \ldots, e_n) \cdot \mathrm{adjoint}(L)(L(x)) \\ & = & \breve{\Delta}_L(e_1, \ldots, e_n)(L(x)) \\ & = & \sum_{i \in \{1, \ldots, n\}} (-1)^{i-1} \cdot \Delta(L(x), L(e_1), \ldots, L(e_{i-1}), \\ & & L(e_{i+1}), \ldots, L(e_n)) \cdot L(e_i) \\ & = & \sum_{i \in \{1, \ldots, n\}} (-1)^{i-1} \cdot \Delta_L(x, e_1, \ldots, e_{i-1}.e_{i+1}, \ldots, e_n) \\ & = & \widehat{\Delta_L}(x, (e_1, \ldots, e_n)) \\ & \underset{10.270}{=} & \Delta_L(e_1, \ldots, e_n) \cdot x \\ & \underset{10.274}{=} & (\det(L) \cdot \Delta(e_1, \ldots, e_n)) \cdot x \\ & = & \det(L) \cdot x \\ & = & (\det(L) \cdot 1_X)(x)\end{array}$$

2. *Let* $(e_1, \ldots, e_n)$ *is a basis of* X *and choose* Δ *such that* $\Delta(e_1, \ldots, e_n) = 1$ *then we have*

$$\begin{array}{rcl}(L \circ \mathrm{adjoint}(L))(t) & = & L(\mathrm{adjoint}(L)(t)) \\ & = & L(\Delta(e_1, \ldots, e_n) \cdot \mathrm{adjoint}(L)(t)) \\ & = & L(\breve{\Delta}_L(e_1, \ldots, e_n)(t)) \\ & = & L\left(\sum_{j \in \{1, \ldots, n\}} (-1)^{j-1} \cdot \Delta(t, L(e_1), \ldots, L(e_{j-1}), \right. \\ & & \left. L(e_{j+1}), \ldots, L(e_n)) \cdot e_j \right) \\ & \underset{L\,\mathrm{is\,linear}}{=} & \sum_{j \in \{1, \ldots, n\}} (-1)^{j-1} \cdot \Delta(t, L(e_1), \ldots, L(e_{j-1}), \\ & & L(e_{j+1}), \ldots, L(e_n)) \cdot L(e_j) \\ & = & \hat{\Delta}_L(t, (L(e_1), \ldots, L(e_n))) \\ & \underset{10.270}{=} & \Delta(L(e_1), \ldots, L(e_n)) \cdot t \\ & = & \Delta_L(e_1, \ldots, e_n) \cdot t \\ & = & (\det(L) \cdot \Delta(e_1, \ldots, e_n)) \cdot t \\ & = & \det(L) \cdot t \\ & = & (\det(L) \cdot 1_X)(t)\end{array}$$

□

Corollary 10.288. *Let* $n \in \mathbb{N}$, X *be a* n*-dimensional vector space over a field* F *of characterization zero,* $L \in \mathrm{Hom}(X, X)$ *then* L *is a isomorphism if and only if* $\det(L) \neq 0$ *and if* L *is a isomorphism then* $L^{-1} = (\det(L))^{-1} \cdot \mathrm{adjoint}(L)$

Proof. Using 10.280 and 10.212 we have that L is a isomorphism if and only if $\det(L) \neq 0$ proving the first part of the corollary. For the second part if L is a isomorphism then $\det(L) \neq 0$ and using the above theorem we have

$$\begin{aligned} \frac{\text{adjoint}(L)}{\det(L)} \circ L &= 1_X \\ L \circ \frac{\text{adjoint}(L)}{\det(L)} &= 1_X \end{aligned}$$

proving that $L^{-1} = \frac{\text{adjoint}(L)}{\det(L)} = (\det(L))^{-1} \cdot \text{adjoint}(L)$ □

10.9 Matrices

Definition 10.289. *Let F be a field, $n, m \in \mathbb{N}$ then a $n \times m$ matrix in F is a mapping $M: \{1, \ldots, n\} \times \{1, \ldots, m\} \rightarrow F$ (so a $n \times m$ matrix is a element of $F^{\{1, \ldots, n\} \times \{1, \ldots, m\}}$. We note $M_{i,j}$ for $M(i, j)$ and call $\{1, \ldots, n\}$ the row indexes of M and $\{1, \ldots, m\}$ the column indexes of M. A $1 \times m$ matrix is called is called a row matrix and a $n \times 1$ matrix is called a column matrix.*

Notation 10.290. *A $n \times m$ matrix M is also noted as*

$$\begin{pmatrix} M_{1,1} & \ldots & M_{1,m} \\ \ldots & \ldots & \ldots \\ M_{n,1} & \ldots & M_{n,m} \end{pmatrix}$$

a row matrix R is noted as

$$\begin{pmatrix} R_{1,1} & \ldots & R_{1,m} \end{pmatrix}$$

and a column matrix C is noted as

$$\begin{pmatrix} C_{1,1} \\ \ldots \\ C_{n,1} \end{pmatrix}$$

***Example* 10.291.** *Given $n \in \mathbb{N}$ then the $n \times n$ matrix E defined by $E: \{1, \ldots, n\} \times \{1, \ldots, n\}$ defined by $(i, j) \rightarrow E(i, j) = E_{i,j} = \delta_{i,j}$ or in matrix notation*

$$\begin{pmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & \ldots & 0 & 1 \end{pmatrix}$$

Definition 10.292. *Let F be a field then we define the sum, scalar product, product as follows*

1. *If A, B are $n \times m$ matrices then $A + B: \{1, \ldots, n\} \times \{1, \ldots, m\} \rightarrow F$ is defined by $(i, j) \rightarrow (A + B)(i, j) = A(i) + B(j)$ or in our shorter notation $(A + B)_{i,j} = A_{i,j} + B_{i,j}$*
2. *If A is a $n \times m$ matrix and $\alpha \in F$ then $\alpha \cdot A: \{1, \ldots, n\} \times \{1, \ldots, m\} \rightarrow F$ is defined by $(i, j) \rightarrow (\alpha \cdot A)(i, j) = \alpha \cdot A(i, j)$ or in our shorter notation $(\alpha \cdot A)_{i,j} = \alpha \cdot A_{i,j}$*

3. *If A is a $n \times m$ matrix and B is a $m \times k$ matrix then $A \cdot B: \{1, ..., n\} \times \{1, ..., k\} \rightarrow F$ is defined by $(i, j) \rightarrow (A \cdot B)(i, j) = \sum_{l=1}^{m} A(i, l) \cdot B(l, j)$ or in our shorter notation $(A \cdot B)_{i,j} = \sum_{l=1}^{m} A_{i,l} \cdot B_{l,j}$*

Definition 10.293. *Let F be a field, $n, m \in \mathbb{N}$ then $\mathcal{M}(n \times m, F) = F^{\{1, ..., n\} \times \{1, ..., m\}}$ (the set of $n \times m$ matrices)*

Theorem 10.294. *Let F be a field, $n, m \in \mathbb{N}$ then $\langle \mathcal{M}(n \times m, F), +, \cdot \rangle$ forms a vector space over F. The zero element is the matrix 0 where $\forall i, j \in \{1, ..., n\} \times \{1, ..., m\}$ we have $0_{i,j} = 0$ which as usually noted by 0.*

Proof. First we prove the group axioms ;

1. **(associativity)** If $A, B, C \in \mathcal{M}(n \times m, F)$ then $((A + B) + C)_{i,j} = (A + B)_{i,j} + C_{i,j} = ((A_{i,j} + B_{i,j}) + C_{i,j}) = A_{i,j} + (B_{i,j} + C_{i,j}) = A_{i,j} + (B + C)_{i,j} = (A + (B + C))_{i,j} \Rightarrow ((A + B) + C) = (A + (B + C))$

2. **(commutativity)** If $A, B \in \mathcal{M}(n \times m, F)$ then $(A + B)_{i,j} = A_{i,j} + B_{i,j} = B_{i,j} + A_{i,j} = (B + A)_{i,j} \Rightarrow A + B = B + A$

3. **(neutral element)** $(A + 0)_{i,j} + A_{i,j} + 0_{i,j} = A_{i,j} + 0 = A_{i.j} \Rightarrow A + 0 = A$ using commutativity we have then $0 + A = a$

4. **(inverse element)** If $A \in \mathcal{M}(n \times m, F)$ define then $-A$ by $\forall (i, j) \in \{1, ..., n\} \times \{1, ..., m\}$ we have $(-A)_{i,j} = -A_{i,j}$ so that $(A + (-A))_{i,j} = A_{i,j} + (-A)_{i,j} = A_{i,j} + (-A_{i,j}) = 0 \Rightarrow A + (-A) = 0$ and using commutativity we have then also $(-A) + A = 0$

For the rest of axioms note:

1. If $A, B \in \mathcal{M}(n \times m, F)$ and $\alpha \in F$ then $(\alpha \cdot (A + B))_{i,j} = \alpha \cdot (A + B)_{i,j} = \alpha \cdot (A_{i,j} + B_{i,j}) = \alpha \cdot A_{i,j} + \alpha \cdot B_{i,j} = (\alpha \cdot A)_{i,j} + (\alpha \cdot B)_{i,j} = (\alpha \cdot A + \alpha \cdot B)_{i,j} \Rightarrow \alpha \cdot (A + B) = \alpha \cdot A + \alpha \cdot B$

2. If $\alpha, \beta \in F$ and $A \in \mathcal{M}(n \times m, F)$ then $((\alpha + \beta) \cdot A)_{i,j} = (\alpha + \beta) \cdot A_{i,j} = \alpha \cdot A_{i,j} + \beta \cdot A_{i,j} = (\alpha \cdot A)_{i,j} + (\beta \cdot A)_{i,j} = (\alpha \cdot A + \beta \cdot A)_{i,j} \Rightarrow (\alpha + \beta) \cdot A = \alpha \cdot A + \beta \cdot A$

3. If $\alpha, \beta \in F$ and $A \in \mathcal{M}(n \times m, F)$ then $((\alpha \cdot \beta) \cdot A)_{i,j} = (\alpha \cdot \beta) \cdot A_{i,j} = \alpha \cdot (\beta \cdot A_{i,j}) = \alpha \cdot (\beta \cdot A)_{i,j} = (\alpha \cdot (\beta \cdot A))_{i,j} \Rightarrow (\alpha \cdot \beta) \cdot A = \alpha \cdot (\beta \cdot A)$

4. If 1 is the neutral element in F, $A \in \mathcal{M}(n \times m, F)$ then $(1 \cdot A)_{i,j} = 1 \cdot A_{i,j} = A_{i,j} \Rightarrow (1 \cdot A) = A$ □

Theorem 10.295. *Given $n \in \mathbb{N}$ and a field F then $\langle \mathcal{M}(n \times n, F), \cdot \rangle$ is a semi-group with muliplicative neutral element E*

Proof.

associativity. *Given $A, B, C \in \mathcal{M}(n \times n, F)$ then for $i, j \in \{1, ..., n\}$ we have*

$$\begin{array}{rcl}((A \cdot B) \cdot C)_{i,j} & = & \displaystyle\sum_{k=1}^{n} (A \cdot B)_{i,k} \cdot C_{k,j} \\ & = & \displaystyle\sum_{k=1}^{n} \left(\sum_{l=1}^{n} A_{i,l} \cdot B_{l,k}\right) \cdot C_{k,j} \\ & = & \displaystyle\sum_{k=1}^{n} \left(\sum_{l=1}^{n} ((A_{i,l} \cdot B_{l,k}) \cdot C_{k,j})\right) \\ & \underset{10.48}{=} & \displaystyle\sum_{l=1}^{n} \left(\sum_{k=1}^{n} (A_{i,l} \cdot B_{l,k}) \cdot C_{k,j}\right) \\ & = & \displaystyle\sum_{l=1}^{n} \left(\sum_{k=1}^{n} A_{i,l} \cdot (B_{l,k} \cdot C_{k,j}\right) \\ & = & \displaystyle\sum_{l=1}^{n} A_{i,l} \cdot \left(\sum_{k=1}^{n} (B_{l,k} \cdot C_{k,j}\right) \\ & = & \displaystyle\sum_{l=1}^{n} A_{i,l} \cdot (B \cdot C)_{l,j} \\ & = & ((A \cdot B) \cdot C)_{i,j}\end{array}$$

proving that $(A \cdot B) \cdot C = A \cdot (B \cdot C)$

neutral element. *Let $A \in \mathcal{M}(n \times n, F)$ and $i, j \in \{1, ..., n\}$ then*

$$\begin{array}{rcl}(A \cdot E)_{i,j} & = & \displaystyle\sum_{k=1}^{n} A_{i,k} \cdot E_{k,j} \\ & = & \displaystyle\sum_{k=1}^{n} A_{i,k} \cdot \delta_{k,j} \\ & = & A_{i,j}\end{array}$$

proving that $A \cdot E = A$ and

$$\begin{array}{rcl}(E \cdot A)_{i,j} & = & \displaystyle\sum_{k=1}^{n} E_{i,k} \cdot A_{k,j} \\ & = & \displaystyle\sum_{k=1}^{n} \delta_{i,k} \cdot A_{k,j} \\ & = & A_{i,j}\end{array}$$

proving that $E \cdot A = A$ □

Definition 10.296. *Let F be a field, $n, m \in \mathbb{N}$ and $M \in \mathcal{M}(n \times m, F)$ then $M^T \in \mathcal{M}(m \times n, F)$ is defined by $(M^T)_{i,j} = M_{j,i}$*

Theorem 10.297. *Let F be a field $n, m \in \mathbb{N}$ and $M, N \in \mathcal{M}(n \times n, F)$ then $(M \cdot N)^T = N^T \cdot M^T$*

Proof. Let $i, j \in \{1, \ldots, n\}$ then

$$\begin{aligned}((M \cdot N)^T)_{i,j} &= (M \cdot N)_{j,i} \\ &= \sum_{k=1}^{n} M_{j,k} \cdot N_{k,i} \\ &= \sum_{k=1}^{n} (M^T)_{k,j} \cdot (N^T)_{i,k} \\ &= (N^T \cdot M^T)_{i,j}\end{aligned}$$

$\square$

Definition 10.298. *Let F be a field, $n, m \in \mathbb{N}$ and $M \in \mathcal{M}(n \times m)$ then we define the following:*

1. $\forall i \in \{1, \ldots, n\}$ $\operatorname{row}(M, i) = (M_{i,1}, \ldots, M_{i,m}) \in F^m$ *(so* $\forall j \in \{1, \ldots, m\}$ *we have* $\operatorname{row}(M, i)_j = M_{i,j}$*)*
2. $\forall i \in \{1, \ldots, m\}$ $\operatorname{col}(M, i) = (M_{1,i}, \ldots, M_{n,i}) \in F^n$ *(so* $\forall j \in \{1, \ldots, n\}$ *we have* $\operatorname{col}(M, i)_j = M_{j,i}$*)*
3. $\operatorname{rows}(M) = \{\operatorname{row}(M, i) | i \in \{1, \ldots, n\}\} \subseteq F^m$
4. $\operatorname{cols}(M) = \{\operatorname{col}(M, i) | i \in \{1, \ldots, m\}\} \subseteq F^n$
5. *The row rank of M is the* $\dim(\mathcal{S}(\operatorname{rows}(M)))$ *[which as* $\dim(F^m) = m$ *(see 10.162) and 10.208) must be finite and lower then m]*
6. *The column rank of M is the* $\dim(\mathcal{S}(\operatorname{cols}(M)))$ *[which as* $\dim(F^n) = n$ *(see 10.162) and 10.208) must be finite and lower then n]*

Theorem 10.299. *Let F be a field, $n, m \in \mathbb{N}$ and $M \in \mathcal{M}(n \times m, F)$ then the column rank and the row rank of M are equal. The column (or row) rank of the matrix is called the rank of the matrix , noted by* $\operatorname{rank}(M)$*. Note that by the previous definition we must have that* $\operatorname{rank}(M) \leqslant \min(n, m)$*.*

Proof. Lets assume that the column rank of M is c and the row rank of M is r then we have

1. As the column rank of M is c we have that $\dim(\mathcal{S}(\operatorname{cols}(M))) = r$ and there exists thus a basis $\{e_i\}_{i \in \{1, \ldots, c\}}$ of $\mathcal{S}(\operatorname{cols}(M))$. This means that $\forall i \in \{1, \ldots, m\}$ there exists a $\{\lambda_{i,j}\}_{j \in \{1, \ldots, c\}}$ such that $\operatorname{col}(M, i) = \sum_{j=1}^{c} \lambda_{i,j} \cdot e_j$. Now $\forall k \in \{1, \ldots, n\}$ and $\forall i \in \{1, \ldots, m\}$ we have $\operatorname{row}(M, k)_i = M_{k,i} = \operatorname{col}(M, i)_k = \left(\sum_{j=1}^{c} \lambda_{i,j} \cdot e_j\right)_k \underset{10.10}{=} \sum_{j=1}^{c} \lambda_{i,j} \cdot (e_j)_k$. If we define now $\forall j \in \{1, \ldots, c\}$ $f_j = (\lambda_{1,j}, \ldots, \lambda_{m,j}) \in F^m$ so that $\forall i \in \{1, \ldots, m\} \vDash (f_j)_i = \lambda_{i,j}$ then we have that $\operatorname{row}(M, k)_i = \sum_{j=1}^{c} (f_j)_i \cdot (e_j)_k = \sum_{j=1}^{c} (e_j)_k \cdot (f_j)_i = \left(\sum_{j=1}^{c} (e_j)_k \cdot f_j\right)_i \Rightarrow \operatorname{row}(M, k) = \sum_{j=1}^{c} (e_j)_k \cdot f_j$ proving that $\forall k \in \{1, \ldots, m\} \vDash \operatorname{row}(M, k) \in \mathcal{S}(\{f_j | j \in \{1, \ldots, c\}\}) \Rightarrow \operatorname{rows}(M) \subseteq \mathcal{S}(\{f_j | j \in \{1, \ldots, c\}\})$. From this and 10.136 we have that $\mathcal{S}(\operatorname{rows}(M)) \subseteq \mathcal{S}(\mathcal{S}(\{f_j | j \in \{1, \ldots, c\}\})) \underset{10.135}{=} \mathcal{S}(\{f_j | j \in \{1, \ldots, c\}\})$. Using 10.156 we have that $\dim(\mathcal{S}(\{f_j : j \in \{1, \ldots, c\}\})) \leqslant c$ and as $\mathcal{S}(\operatorname{rows}(M)) \subseteq \mathcal{S}(\{f_j | j \in \{1, \ldots, c\}\})$ we have by 10.208 that

$$r = \dim(\mathcal{S}(\operatorname{rows}(M))) \leqslant c. \tag{10.92}$$

2. As the row rank of M is r we have that $\dim(\mathcal{S}(\mathrm{rows}(M)))=r$ and there exists thus a basis $\{e_i\}_{i\in\{1,\ldots,r\}}$ of $\mathcal{S}(\mathrm{rows}(M))$. This means that $\forall i\in\{1,\ldots,n\}$ there exists a $\{\lambda_{i,j}\}_{j\in\{1,\ldots,r\}}$ such that $\mathrm{row}(M,i)=\sum_{j=1}^{r}\lambda_{i,j}\cdot e_j$. Now $\forall k\in\{1,\ldots,m\}$ and $\forall i\in\{1,\ldots,n\}$ we have that $\mathrm{col}(M,k)_i=M_{i,k}=\mathrm{row}(M,i)_k=(\sum_{j=1}^{r}\lambda_{i,j}\cdot e_j)_k=\sum_{j=1}^{r}\lambda_{i,j}(e_j)_k$. If we define now $\forall j\in\{1,\ldots,r\}$ $f_j=(\lambda_{1,j},\ldots,\ldots,\lambda_{n,j})\in F^n$ then $\mathrm{col}(M,k)_i=\sum_{j=1}^{r}(f_j)_i\cdot(e_j)_k=\sum_{j=1}^{r}(e_j)_k\cdot(f_j)_i=(\sum_{j=1}^{r}(e_j)_k\cdot f_j)_i\Rightarrow\mathrm{col}(M,k)=\sum_{j=1}^{r}(e_j)_k\cdot f_j$ proving that $\forall k\in\{1,\ldots,m\}\vDash\mathrm{col}(M,k)\in\mathcal{S}(\{f_j|j\in\{1,\ldots,r\}\})\Rightarrow\mathrm{cols}(M)\subseteq\mathcal{S}(\{f_j|j\in\{1,\ldots,r\}\})$. From this and 10.136 we have that $\mathcal{S}(\mathrm{cols}(M))\subseteq\mathcal{S}(\mathcal{S}(\{f_j|j\in\{1,\ldots,r\}\}))\underset{10.135}{=}\mathcal{S}(\{f_j|j\in\{1,\ldots,r\}\})$. Using 10.156 we have that $\dim(\mathcal{S}(\{f_j:j\in\{1,\ldots,r\}\}))\leqslant r$ and as $\mathcal{S}(\mathrm{cols}(M))\subseteq\mathcal{S}(\{f_j|j\in\{1,\ldots,c\}\})$ we have by 10.208 that

$$c=\dim(\mathcal{S}(\mathrm{cols}(M)))\leqslant r \tag{10.93}$$

Using 10.92 and 10.93 we have finally $c=r$ proving the theorem. □

Definition 10.300. *Let X, Y be finite dimensional vector spaces over a field F with $\dim(X)=n$, $\dim(Y)=m$ and let $L\colon X\to Y$ be a linear mapping ($L\in\mathrm{Hom}(X,Y)$). Then if $\{e_i\}_{i\in\{1,\ldots,n\}}$ is a basis for X, $\{f_i\}_{i\in\{1,\ldots,m\}}$ is a basis of Y we have have $\forall i\in\{1,\ldots,n\}$ the existence of a unique family $\{L_{i,j}\}_{j\in\{1,\ldots,m\}}$ such that $L(e_i)=\sum_{j\in\{1,\ldots,m\}}L_{i,j}\cdot f_j$. This defines a matrix $\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,m\}})\in M(m\times n,F)$ by $\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,m\}})_{i,j}=L_{j,i}=(L(e_j))_i$ so that we have $L(e_i)=\sum_{j\in\{1,\ldots,m\}}\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,m\}})_{j,i}\cdot f_j$. If $\{e_i\}_{i\in\{1,\ldots,n\}}$ and $\{f_i\}_{i\in\{1,\ldots,m\}}$ is assumed to be know we use the shorter notation $\mathcal{M}(L)$.*

Definition 10.301. *Let X be a finite dimensional space with base $\{e_i\}_{i\in\{1,\ldots,n\}}$, $L\in\mathrm{Hom}(X)=\mathrm{Hom}(X,X)$ then $\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}})=\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{e_i\}_{i\in\{1,\ldots,n\}})$*

Example 10.302. Let X be a finite dimensional vector space over a field F with $\dim(X)=n$ and $\{e_i\}_{i\in\{1,\ldots,n\}}$ a basis for X then we have the following equivalences for $L\in\mathrm{Hom}(X)$:

1. $L=1_X$
2. $\mathcal{M}(1_X,\{e_i\}_{i\in\{1,\ldots,n\}},\{e_i\}_{i\in\{1,\ldots,n\}})=E$

Proof.

$1\Rightarrow 2$.

$$\begin{aligned}\mathcal{M}(1_X,\{e_i\}_{i\in\{1,\ldots,n\}},\{e_i\}_{i\in\{1,\ldots,n\}})_{i,j} &= 1_X(e_j)_i\\ &= (e_j)_i\\ &\underset{e_i=\sum_{j\in\{1,\ldots,n\}}\delta_{i,j}\cdot e_j}{=} \delta_{j,i}\\ &= \delta_{i,j}\\ &= E_{i,j}\end{aligned}$$

2 ⇒ 1. Let $x \in X \Rightarrow x = \sum_{i=1}^{n} x_i \cdot e_i$ then

$$\begin{aligned} L(x) = L\left(\sum_{i=1}^{n} x_i \cdot e_i\right) &= \sum_{i=1}^{n} x_i \cdot L(e_i) \\ &= \sum_{i=1}^{n} x_i \cdot \left(\sum_{j=1}^{n} E_{i,j} \cdot e_j\right) \\ &= \sum_{i=1}^{n} x_i \cdot \left(\sum_{j=1}^{n} \delta_{i,j} \cdot e_j\right) \\ &= \sum_{i=1}^{n} x_i \cdot e_i \\ &= x = 1_X(x) \end{aligned}$$

proving that $L = 1_X$ □

Theorem 10.303. *Let X, Y be finite dimensional vector spaces over a field F with $\dim(X) = n$, $\dim(Y) = m$, $\{e_i\}_{i\in\{1,\ldots,n\}}$, $\{f_i\}_{i\in\{1,\ldots,m\}}$ basis for X, Y and let $L \in \mathrm{Hom}(X,Y)$ then we have*

1. *If $x = \sum_{i\in\{1,\ldots,n\}} x_i \cdot e_i$ and $L(x) = \sum_{i\in\{1,\ldots,m\}} L(x)_j \cdot f_j$ (see 10.152) then $L(x)_j = \sum_{i\in\{1,\ldots,n\}} \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,n\}})_{j,i} \cdot x_i$*
2. $\mathrm{rank}(L) = \mathrm{rank}(\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,m\}}))$

Proof.

1. If $x = \sum_{i\in\{1,\ldots,n\}} x_i \cdot e_i$ then $L(x) = L\left(\sum_{i\in\{1,\ldots,n\}} x_i \cdot e_i\right) = \sum_{i\in\{1,\ldots,n\}} x_i \cdot L(e_i) = \sum_{i\in\{1,\ldots,n\}} L(e_i) \cdot x_i = \sum_{i\in\{1,\ldots,n\}} \left(\sum_{j\in\{1,\ldots,m\}} \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,m\}})_{j,i} \cdot f_j\right) \cdot x_i = \sum_{i\in\{1,\ldots,n\}} \left(\sum_{j\in\{1,\ldots,m\}} (\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,m\}})_{j,i} \cdot f_j) \cdot x_i\right) \underset{10.48}{=} \sum_{j\in\{1,\ldots,m\}} \left(\sum_{i\in\{1,\ldots,n\}} (\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,m\}})_{j,i} \cdot f_j) \cdot x_i\right) = \sum_{j\in\{1,\ldots,m\}} \left(\sum_{i\in\{1,\ldots,n\}} \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,m\}})_{j,i} \cdot x_i\right) \cdot f_j \underset{L(x)=\sum_{i\in\{1,\ldots,m\}} L(x)_j \cdot f_j \text{ and expansion in a base is unique}}{\Rightarrow} L(x)_j = \sum_{i\in\{1,\ldots,n\}} \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,m\}})_{j,i} \cdot x_i$
2. Let now $M = \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,m\}})$. To prove that $\mathrm{rank}(L) = \mathrm{rank}(M)$ we proceed as follows:
   - If $y \in L(X)$ then there exists a $x = \sum_{i\in\{1,\ldots,n\}} x_i \cdot e_i \in X$ such that $y = L(x)$ giving that $y = L\left(\sum_{i\in\{1,\ldots,n\}} x_i \cdot e_i\right) = \sum_{i\in\{1,\ldots,n\}} x_i \cdot L(e_i) \in \mathcal{S}(\{L(e_i)\}_{i\in\{1,\ldots,n\}})$ proving that $L(X) \subseteq \mathcal{S}(\{L(e_i)\}_{i\in\{1,\ldots,n\}})$. If $y \in \mathcal{S}(\{L(e_i)\}_{i\in\{1,\ldots,n\}})$ then there exists a $\{\alpha_i\}_{i\in I}$ such that $y = \sum_{i\in\{1,\ldots,n\}} \alpha_i \cdot L(e_i) = L\left(\sum_{i\in\{1,\ldots,n\}} \alpha_i \cdot e_i\right) \in L(X)$ proving that $\mathcal{S}(\{L(e_i)\}_{i\in\{1,\ldots,n\}}) \subseteq L(X)$. So we have

$$\mathcal{S}(\{L(e_i)\}_{i\in\{1,\ldots,n\}}) = L(X) \tag{10.94}$$

Using 10.94 and 10.156 there exists a $J \subseteq \{1, \ldots, n\}$ such that $\{L(e_i)\}_{i \in J}$ forms a basis of $L(X)$. Now as $\operatorname{rank}(L) = \dim(L(X))$ we must have that $\#(J) = \operatorname{rank}(L)$ giving the following:

$$\exists \{i_j\}_{j \in \{1, \ldots, \operatorname{rank}(L)\}} \text{ in } \{1, \ldots, n\} \text{ such that } \{L(e_{i_j})\}_{j \in \{1, \ldots, \operatorname{rank}(L)\}} \text{ is a basis of } L(X) \tag{10.95}$$

- Define now $B_c = \{\operatorname{col}(M, i_j) | j \in \{1, \ldots, \operatorname{rank}(L)\}\} \subseteq \operatorname{cols}(M)$ so that by 10.136 we have $\mathcal{S}(B_c) \subseteq \mathcal{S}(\operatorname{cols}(M))$. If $c \in \mathcal{S}(\operatorname{cols}(M)) \Rightarrow \exists \{\lambda_i\}_{i \in \{1, \ldots, n\}}$ such that $c = \sum_{i \in \{1, \ldots, n\}} \lambda_i \cdot \operatorname{col}(M, i)$, so $\forall k \in \{1, \ldots, m\}$ we have $c_k = \left(\sum_{i \in \{1, \ldots, n\}} \lambda_i \cdot \operatorname{col}(M, i)\right)_k \underset{10.10}{=} \sum_{i \in \{1, \ldots, n\}} \lambda_i \cdot \operatorname{col}(M, i)_k = \sum_{i \in \{1, \ldots, n\}} \lambda_i \cdot M_{k,i}$. Take now $y = \sum_{i \in \{1, \ldots, n\}} \lambda_i \cdot L(e_i) \in L(X)$ (see 10.94) then as $\{L(e_{i_j})\}_{j \in \{1, \ldots, \operatorname{rank}(L)\}}$ is a basis of $L(X)$ (see 10.95) we have that $\exists \{\beta_i\}_{i \in \{1, \ldots, \operatorname{rank}(k)\}}$ such that $y = \sum_{l \in \{1, \ldots, \operatorname{rank}(k)\}} \beta_l \cdot L(e_{i_l}) = \sum_{l \in \{1, \ldots, \operatorname{rank}(L)\}} \beta_l \cdot \left(\sum_{j \in \{1, \ldots, m\}} M_{j, i_l} \cdot f_j\right) = \sum_{j \in \{1, \ldots, m\}} \left(\sum_{l \in \{1, \ldots, \operatorname{rank}(L)\}} \beta_l \cdot M_{j, i_l}\right) \cdot f_j$. So $\sum_{j \in \{1, \ldots, m\}} \left(\sum_{l \in \{1, \ldots, \operatorname{rank}(L)\}} \beta_l \cdot M_{j, i_l}\right) \cdot f_j = y \underset{\text{definition of } y}{=} \sum_{i \in \{1, \ldots, n\}} \lambda_i \cdot L(e_i) = \sum_{i \in \{1, \ldots, n\}} \lambda_i \cdot \left(\sum_{j \in \{1, \ldots, m\}} M_{j,i} \cdot f_j\right) = \sum_{j \in \{1, \ldots, m\}} \left(\sum_{i \in \{1, \ldots, n\}} \lambda_i \cdot M_{j,i}\right) \cdot f_j$ and given the uniqueness of the expansion in the base of $\{f_j\}_{j \in \{1, \ldots, m\}}$ we get that $\sum_{l \in \{1, \ldots, \operatorname{rank}(L)\}} \beta_l \cdot M_{j, i_l} = \sum_{i \in \{1, \ldots, n\}} \lambda_i \cdot M_{j,i} = c_i \Rightarrow c_i = \sum_{l \in \{1, \ldots, \operatorname{rank}(L)\}} \beta_l \cdot \operatorname{col}(M, i_l)_i \Rightarrow c = \sum_{l \in \{1, \ldots, \operatorname{rank}(L)\}} \beta_l \cdot \operatorname{col}(M, i_l) \in \mathcal{S}(B_c)$. This proves the following

$$\mathcal{S}(\{\operatorname{col}(M, i_j) | j \in \{1, \ldots, \operatorname{rank}(L)\}\}) = \mathcal{S}(\operatorname{cols}(M)) \tag{10.96}$$

- Let now $\{\lambda_i\}_{i \in \{1, \ldots, \operatorname{rank}(L)\}}$ be such that $0 = \sum_{j \in \{1, \ldots, \operatorname{rank}(L)\}} \lambda_j \cdot \operatorname{col}(M, i_j)$ we have that $\forall k \in \{1, \ldots, m\}$ that $0 = 0_j = \sum_{j \in \{1, \ldots, \operatorname{rank}(L)\}} \lambda_j \cdot \operatorname{col}(M, i_j)_k = \sum_{j \in \{1, \ldots, \operatorname{rank}(L)\}} \lambda_j \cdot M_{k, i_j}$. Take now $y = \sum_{l \in \{1, \ldots, \operatorname{rank}(L)\}} \lambda_l \cdot L(e_{i_l}) \in L(X)$ (by 10.94) then $y = \sum_{l \in \{1, \ldots, \operatorname{rank}(L)\}} \lambda_l \cdot \left(\sum_{j \in \{1, \ldots, m\}} M_{j, i_l} \cdot f_j\right) = \sum_{j \in \{1, \ldots, m\}} \left(\sum_{l \in \{1, \ldots, \operatorname{rank}(L)\}} \lambda_l \cdot M_{j, i_l}\right) \cdot f_j = \sum_{j \in \{1, \ldots, m\}} 0 \cdot f_j = 0$ giving that $0 = y = \sum_{l \in \{1, \ldots, \operatorname{rank}(L)\}} \lambda_l \cdot L(e_{i_l}) \underset{\{L(e_{i_l})\}_{l \in \{1, \ldots, \operatorname{rank}(L)\}} \text{ is a basis}}{\Rightarrow} \forall l \in \{1, \ldots, \operatorname{rank}(L)\}$ we have $\lambda_l = 0$. This proves that

$$\{\operatorname{col}(M, i_j)\}_{j \in \{1, \ldots, \operatorname{rank}(L)\}} \text{ is linear independent} \tag{10.97}$$

Using 10.96 and 10.97 we conclude that $\{\operatorname{col}(M, i_j)\}_{j \in \{1, \ldots, \operatorname{rank}(L)\}}$ forms a basis of $\mathcal{S}(\operatorname{cols}(M))$ so that $\operatorname{rank}(M) = \operatorname{rank}(L)$ □

Theorem 10.304. *Let X, Y be finite dimensionl vector spaces over a field F with dimensions $n, m \in \mathbb{N}$, $\{e_i\}_{i \in \{1, \ldots, n\}}$, $\{f_i\}_{i \in \{1, \ldots, m\}}$ basis of X, Y, $\alpha \in F$ and $L \in \operatorname{Hom}(X, Y)$ then we have $\mathcal{M}(\alpha \cdot L, \{e_i\}_{i \in \{1, \ldots, n\}}, \{f_i\}_{i \in \{1, \ldots, m\}}) = \alpha \cdot \mathcal{M}(L, \{e_i\}_{i \in \{1, \ldots, n\}}, \{f_i\}_{i \in \{1, \ldots, m\}})$*

Proof. $\mathcal{M}(\alpha\cdot L,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,m\}})_{i,j}=(\alpha\cdot L(e_j))_i=\alpha\cdot(L(e_j)_i)=\alpha\cdot\mathcal{M}(L,\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,m\}})_{i,j}$ □

Theorem 10.305. *Let X,Y and Z be vector spaces over a field F with dimensions n,m and r. Let $\{e_i\}_{i\in\{1,\ldots,n\}}$, $\{f_i\}_{i\in\{1,\ldots,m\}}$ and $\{g_i\}_{i\in\{1,\ldots,r\}}$ be the basis of X, Y and Z, $L_1\in\mathrm{Hom}(X,Y)$ and $L_2\in\mathrm{Hom}(Y,Z)$ then $\mathcal{M}(L_2\circ L_1,\{e_i\}_{i\in\{1,\ldots,n\}},\{g_i\}_{i\in\{1,\ldots,r\}})=\mathcal{M}(L_2,\{f_i\}_{i\in\{1,\ldots,m\}},\{g_i\}_{i\in\{1,\ldots,r\}})\cdot\mathcal{M}(L_1,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,m\}})$*

Proof. $\forall i\in\{1,\ldots,n\}$ we have

$$\begin{aligned}
(L_2\circ L_1)(e_i) &= L_2(L_1(e_i))\\
&= L_2\left(\sum_{k\in\{1,\ldots,m\}}\mathcal{M}((L_1,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{0,\ldots,m\}})_{k,i}\cdot f_k\right)\\
&= \sum_{k\in\{1,\ldots,m\}}\mathcal{M}((L_1,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{0,\ldots,m\}})_{k,i}\cdot L_2(f_k)\\
&= \sum_{k\in\{1,\ldots,m\}}\mathcal{M}(L_1,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{0,\ldots,m\}})_{k,i}\cdot\\
&\quad\left(\sum_{l\in\{1,\ldots,r\}}\mathcal{M}(L_2,\{f_i\}_{i\in\{1,\ldots,m\}},\{g_i\}_{i\in\{1,\ldots,r\}})_{l,k}\cdot g_l\right)\\
&= \sum_{l\in\{1,\ldots,r\}}\left(\sum_{k\in\{1,\ldots,m\}}\mathcal{M}(L_1,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{0,\ldots,m\}})_{k,i}\cdot\right.\\
&\quad\left.\mathcal{M}(L_2,\{f_i\}_{i\in\{1,\ldots,m\}},\{g_i\}_{i\in\{1,\ldots,r\}})_{l,k}\right)\cdot g_l\\
&= \sum_{l\in\{1,\ldots,r\}}\left(\sum_{k\in\{1,\ldots,m\}}\mathcal{M}(L_2,\{f_i\}_{i\in\{1,\ldots,m\}},\{g_i\}_{i\in\{1,\ldots,r\}})_{l,k}\cdot\right.\\
&\quad\left.\mathcal{M}(L_1,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{0,\ldots,m\}})_{k,i}\right)\cdot g_l\\
&= \sum_{l\in\{1,\ldots,r\}}(\mathcal{M}(L_2,\{f_i\}_{i\in\{1,\ldots,m\}},\{g_i\}_{i\in\{1,\ldots,r\}})\cdot\mathcal{M}(L_1,\\
&\quad\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{0,\ldots,m\}}))_{l,i}\cdot g_l
\end{aligned}$$

By the definition of $\mathcal{M}$ we have further $\sum_{l\in\{1,\ldots,r\}}\mathcal{M}(L_2\circ L_1,\{e_i\}_{i\in\{1,\ldots,n\}},\{g_i\}_{i\in\{1,\ldots,r\}})_{l,i}\cdot g_l=(L_2\circ L_1)(e_i)=\sum_{l\in\{1,\ldots,r\}}(\mathcal{M}(L_2,\{f_i\}_{i\in\{1,\ldots,m\}},\{g_i\}_{i\in\{1,\ldots,r\}})\cdot\mathcal{M}(L_1,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{0,\ldots,m\}}))_{l,i}\cdot g_l$ and as the expansion in a basis is unique we have finally

$$\begin{aligned}
\mathcal{M}(L_2\circ L_1,\{e_i\}_{i\in\{1,\ldots,n\}},\\
\{g_i\}_{i\in\{1,\ldots,r\}})_{l,i} &= (\mathcal{M}(L_2,\{f_i\}_{i\in\{1,\ldots,m\}},\{g_i\}_{i\in\{1,\ldots,r\}})\cdot\\
&\quad\mathcal{M}(L_1,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{0,\ldots,m\}}))_{l,i}
\end{aligned}$$

proving the theorem. □

Example 10.306. Let $n\in\mathbb{N}$, $\sigma\in P_n$, F a field and $\{\varepsilon_i\}_{i\in\{1,\ldots,n\}}$ the basis for F^n then we have for $\sigma 1_{F^n}\in\operatorname{Hom}(F^n,F^n)$:

1. $\mathcal{M}(\sigma 1_{F^n},\{\varepsilon_i\}_{i\in\{1,\ldots,n\}},\{\varepsilon_i\}_{i\in\{1,\ldots,n\}})_{i,j}=\delta_{i,\sigma(j)}$
2. If $L\in\operatorname{Hom}(F^n,F^m)$ then $\sigma L=L\circ\sigma 1_{F^n}\in\operatorname{Hom}(F^n,F^m)$
3. $\mathcal{M}(\sigma L,\ \{\varepsilon_i\}_{i\in\{1,\ldots,n\}},\ \{\varepsilon_i\}_{i\in\{1,\ldots,m\}},\ \{\varepsilon_i\}_{i\in\{1,\ldots,m\}})_{i,j}=\mathcal{M}(L\circ\sigma 1_{F_n},\ \{\varepsilon_i\}_{i\in\{1,\ldots,n\}},\ \{\varepsilon_i\}_{i\in\{1,\ldots,m\}},\ \{\varepsilon_i\}_{i\in\{1,\ldots,m\}})_{i,j}=\mathcal{M}(\sigma L,\ \{\varepsilon_i\}_{i\in\{1,\ldots,n\}},\ \{\varepsilon_i\}_{i\in\{1,\ldots,m\}},\{\varepsilon_i\}_{i\in\{1,\ldots,m\}})_{i,\sigma(j)}$

Proof. First by 10.222 we have that σi_F and σL are linear. Next we have

1. $(\sigma 1_{F^n}(\varepsilon_i))_j \underset{\text{think of }\varepsilon_i\text{ as a function of }\{1,\ldots,n\}\to F}{=} (1_F(\varepsilon_i\circ\sigma))_j=(\varepsilon_i\circ\sigma)_j=(\varepsilon_i\circ\sigma)(j)=\varepsilon_i(\sigma(j))=(\varepsilon_i)_{\sigma(j)}=\delta_{i,\sigma(j)}$. Using 10.162 we have that $\sigma 1_{F^n}(\varepsilon_i)=\sum_{j\in\{1,\ldots,n\}}(\sigma 1_{F^n}(\varepsilon_i))_j\cdot\varepsilon_j$ giving that $\mathcal{M}(\sigma 1_F,\{\varepsilon_i\}_{i\in\{1,\ldots,n\}},\{\varepsilon_i\}_{i\in\{1,\ldots,n\}})_{i,j}=\delta_{i,\sigma(j)}$
2. If $x\in F^n$ (and thus $x:\{1,\ldots,n\}\to F$) then we have $\sigma L(x)=L(x\circ\sigma)=L(1_{F^n}(x\circ\sigma))=L(\sigma 1_{F^n}(x))=(L\circ\sigma 1_{F^n})(x)$ so $\sigma L=L\circ\sigma 1_{F^n}$
3. Using 10.305 we have that $\mathcal{M}(L\circ\sigma 1_{F^n},\{\varepsilon\}_{i\in\{1,\ldots,n\}},\{\varepsilon_i\}_{1,\ldots,m})_{k,l}=(\mathcal{M}(L,\{\varepsilon_i\}_{i\in\{1,\ldots,n\}},\{\varepsilon_i\}_{i\in\{1,\ldots,m\}})\cdot\mathcal{M}(1_{F^n},\{\varepsilon_i\}_{i\in\{1,\ldots,n\}},\{\varepsilon_i\}_{i\in\{1,\ldots,n\}}))_{k,l}=\sum_{j\in\{1,\ldots,n\}}\mathcal{M}(L,\{\varepsilon_i\}_{i\in\{1,\ldots,n\}},\{\varepsilon_i\}_{i\in\{1,\ldots,m\}})_{k,j}\cdot\mathcal{M}(1_{F^n},\{\varepsilon_i\}_{i\in\{1,\ldots,n\}},\{\varepsilon_i\}_{i\in\{1,\ldots,n\}})_{j,l}=\sum_{j\in\{1,\ldots,n\}}\mathcal{M}(L,\{\varepsilon_i\}_{i\in\{1,\ldots,n\}},\{\varepsilon_i\}_{i\in\{1,\ldots,m\}})_{k,j}\cdot\delta_{j,\sigma(l)}=\mathcal{M}(L,\{\varepsilon_i\}_{i\in\{1,\ldots,n\}},\{\varepsilon_i\}_{i\in\{1,\ldots,m\}})_{k,\sigma(l)}$ □

Theorem 10.307. *Let X, Y be vector spaces over a field F with dimensions n, m and basis $\{e_i\}_{i\in\{1,\ldots,n\}}$, $\{f_i\}_{i\in\{1,\ldots,m\}}$. Then $\mathcal{M}(\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,m\}})$: $\operatorname{Hom}(X,Y)\to M(m\times n,F)$ defined by $L\to\mathcal{M}(\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,n\}})(L)=\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,n\}})$ is a bijection. Furthermore if $M\in M(m\times n,F)$ then $L=\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,n\}})^{-1}(M)$ is defined by $x=\sum_{i\in\{1,\ldots,n\}}x_i\cdot e_i\to L(x)=\sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,n\}}M_{i,j}\cdot x_j\right)\cdot f_i$. In other words we can define a $L\in\operatorname{Hom}(X,Y)$ by specifying how $\forall i\in\{1,\ldots,n\}$ $L(e_i)$ is written in terms of the basis $\{f_i\}_{i\in\{1,\ldots,m\}}$ as this defines $\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,m\}})$.*

Proof.

1. **(injectivity)** Let L_1, $L_2\in\operatorname{Hom}(X,Y)$ with $\mathcal{M}(\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,n\}})(L_1)=\mathcal{M}(\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,n\}})(L_2)$ then we have that $\mathcal{M}(L_1,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,n\}})=\mathcal{M}(L_2,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,n\}})$. If now $x=\sum_{i\in\{1,\ldots,n\}}x_i\cdot e_i$ then $L_1(x)=\sum_{j\in\{1,\ldots,m\}}\left(\sum_{i\in\{1,\ldots,n\}}\mathcal{M}(L_1,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,n\}})_{i,j}\cdot x_j\right)\cdot f_j=\sum_{j\in\{1,\ldots,m\}}\left(\sum_{i\in\{1,\ldots,n\}}\mathcal{M}(L_2,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,n\}})_{i,j}\cdot x_j\right)\cdot f_j=L_2(x)$ proving that $L_1=L_2$
2. **(surjectivity)** Let $M\in M(m\times n,F)$ and define $L: X\to Y$ by $x=\sum_{i\in\{1,\ldots,n\}}x_i\to L(x)=\sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,n\}}M_{i,j}\cdot x_j\right)\cdot f_i$ then we have:
   a. **($L\in\operatorname{Hom}(X,Y)$)** For if $\alpha,\beta\in F$, $x=\sum_{i\in\{1,\ldots,n\}}x_i\cdot e_i\in X$, $y=\sum_{i\in\{1,\ldots,n\}}y_i\cdot e_i\in X$ then $L(\alpha\cdot x+\beta\cdot y)=\sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,n\}}M_{i,j}\cdot(\alpha\cdot x+\beta\cdot y)_j\right)\cdot f_i=$

$\sum_{i\in\{1,\ldots,m\}} \left(\sum_{j\in\{1,\ldots,n\}} M_{i,j}\cdot(\alpha\cdot x_j+\beta\cdot y_j)\right)\cdot f_i = \sum_{i\in\{1,\ldots,m\}} \left(\sum_{j\in\{1,\ldots,n\}} \alpha\cdot(M_{i,j}\cdot x_j)+\beta\cdot(M_{i,j}\cdot y_j)\right)\cdot f_i = \sum_{i\in\{1,\ldots,m\}} \left(\alpha\cdot\sum_{j\in\{1,\ldots,n\}} M_{i,j}\cdot x_j+\beta\cdot\sum_{j\in\{1,\ldots,n\}} M_{i,j}\cdot y_j\right)\cdot f_i=\alpha\cdot \sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,n\}} M_{i,j}\cdot x_j\right)+\beta\cdot\sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,n\}} M_{i,j}\cdot y_j\right)=\alpha\cdot L(x)+\beta\cdot L(y)$

b. $\boldsymbol{(\mathcal{M}(\{e_i\}_{i\in\{1,\ldots,n\}},\ \{f_i\}_{i\in\{1,\ldots,m\}})(L)=M)}$ For by definition $(\mathcal{M}(\{e_i\}_{i\in\{1,\ldots,n\}},\ \{f_i\}_{i\in\{1,\ldots,m\}})(L))_{j,i} = (\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,m\}})(L))_{j,i} = L(e_i)_j = \sum_{k\in\{1,\ldots,n\}} M_{j,k}\cdot (e_i)_k \underset{e_i=\sum_{j\in\{1,\ldots,m\}}\delta_{i,j}\cdot e_j}{=} \sum_{k\in\{1,\ldots,n\}} M_{j,k}\cdot\delta_{i,k}=M_{j,i}$ proving (b).

□

Definition 10.308. *Let $n\in\mathbb{N}$, F a field then $M\in\mathcal{M}(n\times n,F)$ then a $N\in\mathcal{M}(n\times n,F)$ is called a* ***inverse*** *of M if $M\cdot N=E=N\cdot M$*

The next theorem proves that the inverse is of a matrix is unique

Theorem 10.309. *Let $n\in\mathbb{N}$, F a field and assume that $M\in\mathcal{M}(n\times n,F)$ has two inverses N,N' then $N=N'$*

Proof. Let N,N' be two inverses of M then $N=N\cdot E=N\cdot(M\cdot N')=(N\cdot M)\cdot N'=E\cdot N'=N'$ □

Definition 10.310. *Let $n\in\mathbb{N}$, F a field and $M\in\mathcal{M}(n\times n,F)$ then if M has a inverse we note the unique inverse of M by M^{-1}.*

Theorem 10.311. *Let $n\in\mathbb{N}$, F a field and $M\in\mathcal{M}(n\times n,F)$ then if M has a inverse M^{-1} then M^{-1} has also a inverse and $(M^{-1})^{-1}=M$*

Proof. This is trivial as $(M^{-1})\cdot M=E=M\cdot(M^{-1})\cdot M$ so that M is a inverse of M^{-1} hence $(M^{-1})^{-1}=M$. □

Theorem 10.312. *If X,Y are finite n-dimensional vector spaces over a field $\mathcal{F}$ with basis $\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,n\}}$ and $L\in\mathrm{Hom}(X,Y)$ then we have the following equivalences*

1. *L has a inverse L^{-1} (or equivalent L is a isomorphism)*
2. *$\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,n\}})$ has a inverse*

Further if L^{-1} exists then $\mathcal{M}(L^{-1},\{f_i\}_{i\in\{1,\ldots,n\}},\{e_i\}_{i\in\{1,\ldots,n\}})=\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,n\}})^{-1}$

Proof.

1 ⇒ 2. We have

$$\begin{aligned} E &= \mathcal{M}(1_Y,\{f_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,n\}}) \\ &= \mathcal{M}(L\circ L^{-1},\{f_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,n\}}) \\ &\underset{10.305}{=} \mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,n\}})\cdot\mathcal{M}(L^{-1},\{f_i\}_{i\in\{1,\ldots,n\}},\{e_i\}_{i\in\{1,\ldots,n\}}) \end{aligned}$$

and

$$\begin{aligned}E &= \mathcal{M}(1_X, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}}) \\ &= \mathcal{M}(L^{-1}\circ L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}}) \\ &\underset{10.305}{=} \mathcal{M}(L^{-1}, \{f_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}}) \cdot \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,n\}})\end{aligned}$$

proving that $\mathcal{M}(L^{-1}, \{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,n\}}) = \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,n\}})^{-1}$ and thus that $\mathcal{M}(L^{-1}, \{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,n\}})$ has a inverse.

2 ⇒ 1. Define $K \in \mathrm{Hom}(Y, X)$ by $K(f_i) = \sum_{j=1}^{n} \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,n\}})^{-1}_{i,j} \cdot f_j$ so that $\mathcal{M}(K, \{f_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}}) = \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,n\}})^{-1}$. Hence we have that

$$\begin{aligned}\mathcal{M}(K \circ L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}}) &\underset{10.305}{=} \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,n\}})^{-1} \cdot \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,n\}}) \\ &= E\end{aligned}$$

proving (see 10.302) that $K\circ L = 1_X$. Further we have

$$\begin{aligned}\mathcal{M}(L \circ K, \{f_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,n\}}) &\underset{10.305}{=} \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,n\}}) \cdot \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,n\}})^{-1} \\ &= E\end{aligned}$$

proving (see 10.302) that $L\circ K = 1_X$. □

Theorem 10.313. *Let F be a field, $m, n \in \mathbb{N}$, $\{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,m\}}$ the canonical basis of F^n, F^m and $\mathcal{M}(\{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,m\}})\colon \mathrm{Hom}(F^n, F^m) \to M(m\times n, F)$ (defined above) then if $x \in F^n$ we have for $M \in M(m\times n, F)$ that $(\mathcal{M}(\{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,m\}})^{-1}(M)(x))_j = \sum_{i\in\{1,\ldots,n\}} M_{j,i}\cdot x_i$. Or if we identify F^n with $M(n\times 1, F)$ and F^m with $M(m\times 1, F)$ that $\mathcal{M}(\{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,m\}})^{-1}(M)(x) = M\cdot x$*

Proof. If $x \in F^n$ (so $x\colon \{1, \ldots, n\} \to F$) then (see 10.162) $x = \sum_{i\in\{1,\ldots,n\}} x_i \cdot \varepsilon_i$, and by definition (see 10.307) we have $\mathcal{M}(\{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,m\}})^{-1}(M)(x) = \sum_{i\in\{1,\ldots,m\}} \left(\sum_{j\in\{1,\ldots,n\}} M_{i,j}x_j\right) \cdot \varepsilon_i$ so that $(\mathcal{M}(\{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,m\}})^{-1}(M)(x))_k = \left(\sum_{i\in\{1,\ldots,m\}} \left(\sum_{j\in\{1,\ldots,n\}} M_{i,j}x_j\right)\cdot \varepsilon_i\right)_k = \sum_{i\in\{1,\ldots,m\}} \left(\sum_{j\in\{1,\ldots,n\}} M_{i,j}x_j\right)\cdot(\varepsilon_i)_k = \sum_{i\in\{1,\ldots,m\}} \left(\sum_{j\in\{1,\ldots,n\}} M_{i,j}x_j\right)\cdot \delta_{i,k} = \sum_{j\in\{1,\ldots,n\}} M_{k,j}\cdot x_j$ proving the theorem. □

Definition 10.314. *If $n\in\mathbb{N}$, F a field and $M\in M(n\times n,F)$ then* $\det(M)=\sum_{\sigma\in P_n}\operatorname{sign}(\sigma)\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{\sigma(i),i}\right)$. *Another notation that we use is* $|M|=\det(M)$

Example 10.315. If $n=1$ then then $M=(M_{1,1})$ and $P_n=\{1_{\{1\}}\}$ and $|M|=\sum_{\sigma\in\{1_{\{1\}}\}}\left(\prod_{i\in\{1\}}M_{i,\sigma(i)}\right)=M_{1,1}$

Example 10.316. If $n\in\mathbb{N}$, F a field and $M\in M(n\times n,F)$ a diagonal matrix meaning that $\forall i,j\in\{1,\ldots,n\}$ we have $M_{i,j}=\delta_{i,j}\cdot M_{i,i}$ then $\det(M)=\prod_{i=1}^n M_{i,i}$

Proof. First if $\sigma\in P_n$ is such that $\sigma\neq 1_{\{1,\ldots,n\}}$ then $\exists i_0\in\{1,\ldots,n\}$ such that $\sigma(i_0)\neq i+0$ hence $\prod_{i\in\{1,\ldots,n\}}M_{\sigma(i),i}=\left(\prod_{i\in\{1,\ldots,\}\setminus\{i_0\}}M_{\sigma(i_0),i+0}\right)\cdot M_{\sigma(i_0),i+0}=0$. So $\det(M)=\sum_{\sigma\in\mathcal{P}_n}\left(\prod_{i\in\{1,\ldots,n\}}M_{\sigma(i),u}\right)=\left(\sum_{\sigma\in P_n\setminus\{1_{\{1,\ldots,n\}}\}}\left(\prod_{i\in\{1,\ldots,n\}}M_{\sigma(i),i}\right)\right)+\prod_{i\in\{1,\ldots,n\}}M_{1_{\{1,\ldots,n\}}(i),i}=0+\prod_{i\in\{1,\ldots,n\}}M_{i,i}=\prod_{i\in\{1,\ldots,n\}}M_{i,i}$ □

Theorem 10.317. *Let F be a field, $n\in\mathbb{N}$ and $M\in M(n\times n,F)$ then* $\det(M)=\sum_{\sigma\in P_n}\operatorname{sign}(\sigma)\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{i,\sigma(i)}\right)$

Proof. First note that from the commutativity of multiplication we have

$$\begin{aligned}\prod_{i\in\{1,\ldots,n\}}M_{\sigma(i),i} &\underset{10.19\,\text{and}\,\sigma\,\text{is}\,a\,\text{bijection}}{=} \prod_{i\in\{1,\ldots,n\}}M_{\sigma(\sigma^{-1}(i)),\sigma^{-1}(i)}\\ &= \prod_{i\in\{1,\ldots,n\}}M_{i,\sigma^{-1}(i)}\end{aligned}$$

from which it follows that

$$\begin{aligned}\det(M) &= \sum_{\sigma\in P_n}\operatorname{sign}(\sigma)\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{i,\sigma^{-1}(i)}\right)\\ &\underset{10.92}{=} \sum_{\sigma\in P_n}\operatorname{sign}(\sigma^{-1})\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{i,\sigma^{-1}(i)}\right)\\ &\underset{^{-1}\,\text{is}\,a\,\text{bijection}\,\text{from}\,P_n\to P_n\,\text{and}\,10.44}{=} \sum_{\tau\in P_n}\operatorname{sign}(\tau)\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{i,\tau(i)}\right)\end{aligned}$$

□

Corollary 10.318. *Let F be a field, $n\in\mathbb{N}$ and $M\in M(n\times n,F)$ then* $\det(M)=\det(M^T)$

Proof. $\det(M)=\sum_{\sigma\in P_n}\operatorname{sign}(\sigma)\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{\sigma(i),i}\right)\underset{\text{previous theorem}}{=}\sum_{\sigma\in P_n}\operatorname{sign}(\sigma)\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{i,\sigma(i)}\right)=\sum_{\sigma\in P_n}\operatorname{sign}(\sigma)\cdot(M^T)_{\sigma(i),i}=\det(M^T)$ □

Theorem 10.319. *If X is a n-dimensional space over a field F of characterization 0 with basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ then if $L\in\operatorname{Hom}(X,X)$ we have* $\det(L)=\det(\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{e_i\}_{i\in\{1,\ldots,n\}}))$

Proof. Let Δ be a determinant function such that $\Delta(e_1, \ldots, e_n) = 1$ (see 10.263) then we have $\Delta_L(e_1, \ldots, e_n) \underset{\text{definition}\,\Delta_L}{=} \Delta(L(e_1), \ldots, L(e_n)) \underset{10.274}{=} \det(L) \cdot \Delta(e_1, \ldots, e_n) = \det(L)$ so we have

$$\Delta_L(e_1, \ldots, e_n) = \det(L) \tag{10.98}$$

As a determinant function is a skew symmetric linear mapping from $X^n \rightarrow F$ we have (using $M = \mathcal{M}(L, \{e_i\}_{i \in \{1, \ldots, n\}}, \{e_i\}_{i \in \{1, \ldots, n\}})$)

$$\begin{array}{lcl}
\Delta_L(e_1, \ldots, e_n) & = & \Delta(L(e_1), \ldots, L(e_n)) \\
 & = & \Delta\left(\sum_{i \in \{1, \ldots, n\}} M_{i,1} \cdot e_i, \ldots, \sum_{i \in \{1, \ldots, n\}} M_{i,n} \cdot e_i\right) \\
 & \underset{10.260}{=} & \sum_{\sigma \in P_n} \left(\left(\prod_{i \in \{1, \ldots, n\}} M_{\sigma(i),i}\right) \cdot \sigma\Delta(e_1, \ldots, e_n)\right) \\
 & \underset{\Delta\,\text{is}\,\text{skew}\,\text{symmetric}}{=} & \left(\sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot \prod_{i \in \{1, \ldots, n\}} M_{\sigma(i),i}\right) \cdot \Delta(e_1, \ldots, e_n) \\
 & = & \det(M) \cdot \Delta(e_1, \ldots, e_n) \\
 & = & \det(M)
\end{array}$$

Using 10.98 we have finally $\det(L) = \det(M)$ $\square$

Theorem 10.320. *If F is a field of characterization zero, $n \in \mathbb{N}$ and $M_1, M_2 \in M(n \times n, F)$ then*

$$\begin{array}{rcl}
\det(M_1 \cdot M_2) & = & \det(M_1) \cdot \det(M_2) \\
\det(E) & = & 1
\end{array}$$

Proof. Let F^n be the canonical vector space over F with the canonical basis $\{\varepsilon_i\}_{i \in \{1, \ldots, n\}}$ and let $\mathcal{M} = \mathcal{M}(\{\varepsilon_i\}_{i \in \{1, \ldots, n\}}, \{\varepsilon_i\}_{i \in \{1, \ldots, n\}})\colon \text{Hom}(F^n, F^n) \rightarrow M(n \times n, F)$ be the bijection from 10.307. Take then $L_1 = \mathcal{M}^{-1}(M_1)$, $L_2 = \mathcal{M}^{-1}(M_2)$ then we have as $\mathcal{M}(L_1 \circ L_2) \underset{10.305}{=} \mathcal{M}(L_1) \cdot \mathcal{M}(L_2) = \mathcal{M}(\mathcal{M}^{-1}(M_1)) \cdot \mathcal{M}(\mathcal{M}^{-1}(M_2)) = M_1 \cdot M_2$

$$\begin{array}{rcl}
\det(M_1 \cdot M_2) & = & \det(\mathcal{M}(L_1 \circ L_2)) \\
 & \underset{10.319}{=} & \det(L_1 \circ L_2) \\
 & \underset{10.280}{=} & \det(L_1) \cdot \det(L_2) \\
 & \underset{10.319}{=} & \det(M_1) \cdot \det(M_2)
\end{array}$$

For the last part $\mathcal{M}(1_{F^n})_{i,j} = (1_{F^n}(\varepsilon_j))_i = (\varepsilon_j)_i = \delta_{i,j} = E_{i,j}$ so that we have

$$\begin{array}{rcl}
\det(E) & = & \det(\mathcal{M}(1_{F^n})) \\
 & \underset{10.319}{=} & \det(1_{F^n}) \\
 & \underset{10.275}{=} & 1
\end{array}$$

$\square$

Theorem 10.321. *Let $n \in \mathbb{N}$, F a field of characterization zero, $M \in M(n \times n, F)$ and $\{\varepsilon_i\}_{i \in \{1,\ldots,n\}}$ be the canonical basis of F^n then if Δ is a determinant function in F^n such that $\Delta(\varepsilon_1, \ldots, \varepsilon_n) = 1$ then* $\det(M) = \Delta(\mathrm{col}(M,1), \ldots, \mathrm{col}(M, n)) = \Delta(\mathrm{row}(M,1), \ldots, \mathrm{row}(M,n))$

Proof. Take $\mathcal{M} = \mathcal{M}(\{\varepsilon_i\}_{i \in \{1,\ldots,n\}}, \{\varepsilon_i\}_{i \in \{1,\ldots,n\}}): \mathrm{Hom}(F^n, F^n) \to M(n \times n, F)$ be the bijection from 10.307. If $M \in M(n \times n, F)$ define then $L = \mathcal{M}^{-1}(M)$ then we have by 10.313 that $(L(\varepsilon_i))_j = \sum_{k \in \{1,\ldots,n\}} M_{j,k} \cdot (\varepsilon_i)_k = \sum_{k \in \{1,\ldots,n\}} M_{j,k} \cdot \delta_{i,k} = M_{j,i} = \mathrm{col}(M,i)_j$ so that

$$L(\varepsilon_i) = \mathrm{col}(M,i) \tag{10.99}$$

So we have

$$\begin{aligned}
\Delta(\mathrm{col}(M,1), \ldots, \mathrm{col}(M,n)) &= \Delta(L(\varepsilon_1), \ldots., L(\varepsilon_n)) \\
&= \Delta_L(\varepsilon_1, \ldots, \varepsilon_n) \\
&= \det(L) \cdot \Delta(\varepsilon_1, \ldots, \varepsilon_n) \\
&= \det(L) \\
&= \det(\mathcal{M}(L)) \\
&= \det(\mathcal{M}(\mathcal{M}^{-1}(L))) \\
&= \det(M)
\end{aligned}$$

This proves that $\det(M) = \Delta(\mathrm{col}(M,1), \ldots, \mathrm{col}(M,n))$. For the rest note that $\mathrm{col}(M^T, i)_j = (M^T)_{j,i} = M_{i,j} = \mathrm{row}(M,i)_j$ giving

$$\begin{aligned}
\Delta(\mathrm{row}(M,1), \ldots, \mathrm{row}(M,n)) &= \Delta(\mathrm{col}(M^T,1), \ldots \mathrm{col}(M^T,n)) \\
&= \det(M^T) \\
&\underset{10.318}{=} \det(M)
\end{aligned}$$

□

Definition 10.322. *If F is a field of characterization zero, $M \in M(n \times n, F)$ and $\sigma: \{1, \ldots, n\} \to \{1, \ldots, n\}$ then we define*

1. *$M_\sigma \in M(n \times n, F)$ by $\forall k, l \in \{1, \ldots, n\}$ $(M_\sigma)_{k,l} = M_{\sigma(k),l}$*
2. *$M^\sigma \in M(n \times n, F)$ by $\forall k, l \in \{1, \ldots, n\}$ $(M^\sigma)_{k,l} = M_{k,\sigma(l)}$*

Essential the original matrices with rows or columns permuted.

Corollary 10.323. *If F is a field of characterization zero, $n \in \mathbb{N}$ and $M \in M(n \times n, F)$ then if $\sigma: \{1, \ldots, n\} \to \{1, \ldots, n\}$ is a permutation then we have* $\det(M_\sigma) = \mathrm{sign}(\sigma) \cdot \det(M)$ *and* $\det(M^\sigma) = \mathrm{sign}(\sigma) \cdot \det(M)$

Proof. $\forall i \in \{1, \ldots, n\}$ we have $\mathrm{row}(M_\sigma, i)_j = (M_\sigma)_{i,j} = M_{\sigma(i),j} = \mathrm{row}(M, \sigma(i))$, then $\det(M_\sigma) = \Delta(\mathrm{row}(M, \sigma(1)), \ldots, \mathrm{row}(M, \sigma(n))) = \mathrm{sign}(\sigma) \cdot \Delta(\mathrm{row}(M,1), \ldots, \mathrm{row}(M, n)) = \mathrm{sign}(\sigma) \cdot \det(M)$.

$\forall i \in \{1, ..., n\}$ we have $\mathrm{col}(M^\sigma, i)_j = (M^\sigma)_{j,i} = M_{j,\sigma(i)} = \mathrm{col}(M, \sigma(i))$, then $\det(M^\sigma) = \Delta(\mathrm{col}(M^\sigma, 1), ..., \mathrm{col}(M^\sigma, n)) = \mathrm{sign}(\sigma) \cdot \Delta(\mathrm{col}(M, 1), ..., \mathrm{col}(M, n)) = \mathrm{sign}(\sigma) \cdot \det(M)$. □

Corollary 10.324. *Let $n \in \mathbb{N}$, F a field of characterization zero then if $M \in M(n \times n, F)$ we have*

$$\det(M) = 0 \Leftrightarrow \mathrm{rank}(M) < n$$

Proof. This follows from $\det(M) = \Delta(\mathrm{row}(M,1), ..., \mathrm{row}(M,n)) = \Delta(\mathrm{col}(M,1), ..., \mathrm{col}(M,n))$ and 10.257 (skew symmetric mappings an linear Independence) and the definition of the rank of a matrix (see 10.299). □

Definition 10.325. *Let $n \in \mathbb{N}$, X a set, $a \in X$ and $x \in X^n$ define $[+a]x \in X^{n+1}$ by $\forall i \in \{1, ..., n+1\}$ $([+a]x)_i = \begin{cases} a \text{ if } i = 1 \\ x_{i-1} \text{ if } i > 1 \end{cases}$. If $x = (x_1, ..., x_n)$ then we sometimes write $[+a]x$ as $(a, x_1, ..., x_n)$*

Lemma 10.326. *Let $n \in \mathbb{N}$, F a field and $0 \in F$ the neutral element in F, then $[+0]: F^n \to F^{n+1}$ defined by $x \to [+0]x$ is a linear mapping.*

Proof. Let $\alpha, \beta \in F$ and $x, y \in F^n$ then for $i \in \{1, ..., n\}$ we have $([+0](\alpha \cdot x + \beta \cdot y))_i = \begin{cases} 0 \text{ if } i = 1 \\ (\alpha \cdot x + \beta \cdot y)_{i-1} \text{ if } i > 1 \end{cases} = \begin{cases} \alpha \cdot 0 + \beta \cdot 0 \text{ if } i = 1 \\ \alpha \cdot x_{i-1} + \beta \cdot y_{i-1} \text{ if } i > 1 \end{cases} = \alpha \cdot [+0]x + \beta \cdot [+0y]$ □

Lemma 10.327. *Let $n \in \mathbb{N} \backslash \{1\}$, F a field of characterization zero, $\{\varepsilon_i\}_{i \in \{1, ..., n\}}$ the canonical basis of F^n then if $\Delta: (F^n)^n \to F$ is a determinant function then $\Delta_{-1}: (F^{n-1})^{n-1} \to F$ defined by $(a_1, ..., a_{n-1}) \to \Delta_{-1}(a_1, ..., a_{n-1}) = \Delta(\varepsilon_1, [+0]a_1, ..., [+0]a_{n-1})$ is a determinant function. Further if $\Delta(\varepsilon_1, ..., \varepsilon_n) = 1$ then $\Delta_{-1}(\varepsilon_1, ..., \varepsilon_{n-1}) = 1$*

Proof. We prove first that Δ_{-1} is a determinant function

1. **(multilinearity)** If $\alpha, \beta \in F$, $x, y \in F^{n-1}$, $i \in \{1, ., n-1\}$ and $a = (a_1, ..., a_{i-1}, \alpha \cdot x + \beta \cdot y, a_{i+1}, ..., a_{n-1}) \in (F^{n-1})^{n-1}$ then $\Delta_{-1}(a_1, ..., a_{i-1}, \alpha \cdot x + \beta \cdot y, a_{i+1}, ..., a_{n-1}) = \Delta(\varepsilon_1, [+0]a_1, ..., [+0]a_{i-1}, [+0](\alpha \cdot x + \beta \cdot y), [+0]a_{i+1}, ..., [+0]a_{n-1}) \underset{\text{previous lemma}}{=} \Delta(\varepsilon_1, [+0]a_1, ..., [+0]a_{i-1}, \alpha \cdot [+0]x + \beta \cdot [+0]y, [+0]a_{i+1}, ..., [+0]a_{n-1}) \underset{\Delta \text{ is multilinear}}{=} \alpha \cdot \Delta(\varepsilon_1, [+0]a_1, ..., [+]a_{i-1}, [+0]x, [+]a_{i+1}, ..., [+0]a_{n-1}) + \beta \cdot (\varepsilon_1, [+0]a_1, ..., [+]a_{i-1}, [+0]y, [+]a_{i+1}, ..., [+0]a_{n-1}) = \alpha \cdot \Delta_{-1}(a_1, ..., a_{i-1}, x, a_{i+1}, ..., a_{n-1}) + \beta \cdot \Delta_{-1}(a_1, ..., a_{i-1}, y, a_{i+1}, ..., a_{n-1})$

2. **(skew symmetry)** If $\{a_1, ..., a_{n-1}\}$ is such that there exists a $i \neq j$ with $a_i = a_j \Rightarrow [+0]a_i = [+0]a_j$ so that $\Delta_{-1}(a_1, ..., a_{n-1}) = \Delta(\varepsilon_1, [+0]a_1, ..., [+0]a_{n-1}) \underset{10.257}{=} 0$ which by 10.257 again means that Δ_{-1} is skew symmetric.

Next if $\Delta(\varepsilon_1, \ldots, \varepsilon_n) = 1$ and as for $i, j \in \{1, \ldots, n-1\}$ we have $([+0]\varepsilon_i)_j = \left\{ \begin{smallmatrix} 0 \text{ if } j=1 \\ (\varepsilon_i)_{j-1} \text{ if } 1<j \end{smallmatrix} \right. = \left\{ \begin{smallmatrix} 0 \text{ if } j=1 \\ \delta_{i,j-1 \text{ if } j>0} \end{smallmatrix} \right. = \delta_{i+1,j} = (\varepsilon_{i+1})_j$ so that $[+0]\varepsilon_i = \varepsilon_{i+1}$ and thus $\Delta_{-1}(\varepsilon_1, \ldots, \varepsilon_{n-1}) = \Delta(\varepsilon_1, [+0]\varepsilon_1, \ldots, [+0]\varepsilon_{n-1}) = \Delta(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) = 1$ □

Definition 10.328. *Let F be a field, $n \in \mathbb{N} \setminus \{1\}$ then we define for $i, j \in \{1, \ldots, n\}$ $[i \boxplus j]: M(n \times n, F) \rightarrow M(n-1 \times n-1, F)$ by $M \rightarrow [i \boxplus j]M$ where*

$$\begin{aligned} ([i \boxplus j]M)_{k,l} &= M_{k,l} \text{ if } k<i \wedge l<j \\ &= M_{k+1,l} \text{ if } i \leqslant k \leqslant n-1 \wedge l<j \\ &= M_{k,l+1} \text{ if } k<i \wedge j \leqslant l \leqslant n-1 \\ &= M_{k+1,l+1} \text{ if } i \leqslant k \leqslant n-1 \wedge j \leqslant l \leqslant n-1 \end{aligned}$$

Lemma 10.329. *Let F be a field, $n \in \mathbb{N} \setminus \{1\}$ then for $i, j \in \{1, \ldots, n\}$ and $M \in M(n \times n, F)$ we have $[i \boxplus j]M^T = ([j \boxplus i]M)^T$*

Proof. We have for $k, l \in \{1, \ldots, n-1\}$ the following cases to consider

$$\begin{aligned} k<i, l<k \Rightarrow ([i \boxplus j]M^T)_{k,l} &= M^T_{k,l} \\ &= M_{l,k} \\ &= ([j \boxplus i]M)_{l,k} \\ &= ([j \boxplus i]M)^T_{k,l} \\ i \leqslant k \leqslant n-1, l<j \Rightarrow ([i \boxplus j]M^T)_{k,l} &= M^T_{k+1,l} \\ &= M_{l,k+1} \\ &= ([j \boxplus i]M)_{l,k} \\ &= ([j \boxplus i]M)^T_{k,l} \\ k<i, j \leqslant l \leqslant n-1 \Rightarrow ([i \boxplus j]M^T)_{k,l} &= M^T_{k,l+1} \\ &= M_{l+1,k} \\ &= ([j \boxplus i]M)_{l,k} \\ &= ([j \boxplus i]M)^T_{k,l} \\ i \leqslant k \leqslant n-1, j \leqslant l \leqslant n-1 \Rightarrow ([i \boxplus j]M^T)_{k,l} &= M^T_{k+1,l+1} \\ &= M_{l+1,k+1} \\ &= ([j \boxplus i]M)_{l,k} \\ &= ([j \boxplus i]M)^T_{k,l} \end{aligned}$$

□

Lemma 10.330. *Let $n \in \mathbb{N} \setminus \{1\}$ and let $M \in M(n \times n, F)$ such that $\forall i \in \{1, \ldots, n\}$ we have $\mathrm{row}(M,1)_i = \delta_{i,1} = \mathrm{col}(M,1)_i$ then $\det(M) = \det([1 \boxplus 1M])$*

Proof. Consider now the following cases for $i \in \{1, \ldots, n\}$ and $\mathrm{row}(M, i)$

1. **($\boldsymbol{i=1}$)** then $\forall j \in \{1, \ldots, n\} \vDash \mathrm{row}(M, i)_j = \delta_{i,1} = (\varepsilon_1)_i$

2. **($i \neq 1$)** then if $j=1$ we have $\operatorname{row}(M,i)_j=\operatorname{row}(M,i)_1=\operatorname{col}(M,1)_i=\delta_{i,1}\underset{i\neq 1}{=}0$ and if $j\in\{2,\ldots,n\}$ we have $\operatorname{row}(M,i)_j=M_{i,j}=([1\boxplus 1]M)_{i-1,j-1}$ so that we have that $\operatorname{row}(M,i)=[+0]\operatorname{row}([1\boxplus 1]M,i-1)$. Let now Δ be a determinant function in F^n such that $\Delta(\varepsilon_1,\ldots,\varepsilon_n)$ then we have by 10.321 that :

$$\begin{array}{lcl}\det(M) & = & \Delta(\operatorname{row}(M,1),\ldots,\operatorname{row}(M,n))\\ & = & \Delta(\varepsilon_1,\operatorname{row}(M,2),\ldots,\operatorname{row}(M,n))\\ & = & \Delta(\varepsilon_1,\operatorname{row}([1\boxplus 1]M,1),\ldots,\operatorname{row}(1\boxplus 1M,n-1))\\ & \underset{10.327}{=} & \Delta(\operatorname{row}([1\boxplus 1]M),\ldots,\operatorname{row}(1\boxplus 1M,n-1))\\ & \underset{10.327\,\text{and}\,10.321}{=} & \det(1\boxplus 1M)\end{array}$$

□

Lemma 10.331. *Let $n\in\mathbb{N}\setminus\{1\}$ and let $M\in M(n\times n,F)$ such that $\exists j\in\{1,\ldots,n\}$ such that $\forall i\in\{1,\ldots,n\}$ we have $\operatorname{row}(M,1)_i=\delta_{j,i}$ and $\operatorname{col}(M,j)_i=\delta_{i,1}$ then $\det(M)=(-1)^{1+j}\cdot\det([1\boxplus j]M)$*

Proof. We have the following cases to consider for for $j\in\{1,\ldots,m\}$:

1. **($j=1$)** then $\forall i\in\{1,\ldots,n\}$ we have $\operatorname{col}(M,j)_i=\delta_{i,1}=\operatorname{row}(M,1)$ so we can use the previous lemma and find that $\det(M)=\det([[1\boxplus 1]M])=(-1)^{1+1}\det([1\boxplus 1]M)\underset{j=1}{=}(-1)^{1+j}\cdot\det([1\boxplus j]M)$

2. **($j\in\{2,\ldots,n\}$)** define then M' by $\forall k,l\in\{1,\ldots,n\}$ $M'_{k,l}=\operatorname{col}(M',l)_k=\operatorname{col}(M,j\rightsquigarrow_n 1(l))_k=M_{,j\rightsquigarrow_n 1(l)}$ then we have $\forall k\in\{1,\ldots,n\}$ that

$$\begin{array}{rcl}\operatorname{row}(M',1)_k & = & \operatorname{col}(M',k)_1\\ & = & \operatorname{col}(M,j\rightsquigarrow_n 1(k))_1\\ & = & \begin{cases}\operatorname{col}(M,j)_1\text{ if }k=1\\ \operatorname{col}(M,k-1)_1\text{ if }1<k\leqslant j\\ \operatorname{col}(M,k)_1\text{ if }j<k\end{cases}\\ & = & \begin{cases}\delta_{1,1}\text{ if }k=1\\ \operatorname{row}(M,1)_{k-1}\text{ if }1<k\leqslant j\\ \operatorname{row}(M,1)_k\text{ if }j<k\end{cases}\\ & = & \begin{cases}1\text{ if }k=1\\ \delta_{j,k-1}\text{ if }1<k\leqslant j\Rightarrow k-1\neq j\\ \delta_{j,k\,\text{if}\,j<k}\end{cases}\\ & = & \begin{cases}1\text{ if }k=1\\ 0\text{ if }1<k\leqslant j\\ 0\text{ if }j<k\end{cases}\\ & = & \delta_{1,k}\end{array}$$

and we have also

$$\begin{array}{rcl}\mathrm{col}(M',1)_k & = & \mathrm{col}(M,j\rightsquigarrow_n 1(1))_k\\ & = & \mathrm{col}(M,j)_k\\ & = & \delta_{1,k}\end{array}$$

So we have $\forall k\in\{1,\ldots,n\}$ that $\mathrm{row}(M',1)_k=\delta_{1,k}=\mathrm{col}(M',1)_k$ so by the previous lemma (see 10.330) we have

$$\det(M')=\det([1\boxplus 1]M') \tag{10.100}$$

Now for $k,l\in\{1,\ldots,n-1\}$ we have as $1\leqslant k,l\leqslant n-1$ that $([1\boxplus 1]M')_{k,l}=M'_{k+1,l+1}$ and we must look at the following possible cases for l where $1\leqslant k\leqslant n-1$

a. $\boldsymbol{(1\leqslant k\leqslant n-1\wedge l<j)}$ then

$$\begin{array}{rcl}M'_{k+1,l+1} & = & \mathrm{col}(M',l+1)_{k+1}\\ & = & \mathrm{col}(M',j\rightsquigarrow_n 1(l+1))_{k+1}\\ & \underset{l<j\Rightarrow 1<l+1\leqslant j}{=} & \mathrm{col}(M,l)_{k+1}\\ & = & M_{k+1,l}\\ & \underset{1\leqslant k<k+1,l<j}{=} & ([1\boxplus j]M)_{k,l}\end{array}$$

b. $\boldsymbol{(1\leqslant k\leqslant n-1\wedge j\leqslant l\leqslant n-1)}$ then

$$\begin{array}{rcl}M'_{k+1,l+1} & = & \mathrm{col}(M',l+1)_{k+1}\\ & = & \mathrm{col}(M,j\rightsquigarrow_n 1(l+1))_{k+1}\\ & \underset{j\leqslant l\Rightarrow j<l+1}{=} & \mathrm{col}(M,l+1)_{k+1}\\ & = & M_{k+1,l+1}\\ & \underset{1\leqslant k<k+1,j<l+1}{=} & ([1\boxplus j]M)_{k,l}\end{array}$$

The above means that $[1\boxplus 1]M'=[1\boxplus j]M$ and given 10.100 we have then

$$\det(M')=\det([1\boxplus j]M) \tag{10.101}$$

Now take a determinant function Δ in F^n such that $\Delta(\varepsilon_1,\ldots,\varepsilon_n)=1$ then by 10.321 we have

$$\begin{array}{rcl}\det(M') & = & \Delta(\mathrm{col}(M',1),\ldots,\mathrm{col}(M',n))\\ & = & \Delta(\mathrm{col}(M,j\rightsquigarrow_n 1(1)),\ldots,\mathrm{col}(M,j\rightsquigarrow_n 1(n)))\\ & \underset{10.96}{=} & (-1)^{j-1}\Delta(\mathrm{col}(M,1),\ldots,\mathrm{col}(M,n))\\ & \underset{10.321}{=} & (-1)^{j-1}\det(M)\end{array}$$

as $\frac{1}{(-1)^{j-1}} = (-1)^{j-1} = (-1)^2 \cdot (-1)^{j-1} = (-1)^{j+1}$ we have $(-1)^{j+1} \cdot \det(M') = \det(M)$ and using 10.101 this gives finally $\det(M) = (-1)^{j+1} \cdot \det([1 \boxplus j]M)$ ☐

Theorem 10.332. *Let $n \in \mathbb{N} \setminus \{1\}$ and for $M \in M(n \times n, F)$ there exists a i, $j \in \{1, \ldots, n\}$ such that $\mathrm{row}(M, i)_k = \delta_{j,k}$ and $\mathrm{col}(M, j)_k = \delta_{i,k}$ then $\det(M) = (-1)^{i+j} \det([i \boxplus j]M)$*

Proof. We have the following cases to consider for $i \in \{1, \ldots, n\}$

1. $\boldsymbol{(i = 1)}$ then $\forall k \in \{1, \ldots, n\}$ we have $\mathrm{row}(M, 1)_k = \delta_{j,k}$ and $\mathrm{col}(M, j)_k = \delta_{1,k}$ so using the previous lemma 10.331 we have that $\det(M) = (-1)^{j+1} \det([1 \boxplus j]M) \underset{i=1}{=} (-1)^{i+j} \det([i \boxplus j]M)$

2. $\boldsymbol{(i \in \{2, \ldots, n\})}$ define then M' by $\forall k, l \in \{1, \ldots, n\}$ $M'_{k,l} = \mathrm{row}(M', k)_l = \mathrm{row}(M, i \rightsquigarrow_n 1(k))_l = M_{i \rightsquigarrow_n 1(k), l}$ then we have $\forall k \in \{1, \ldots, n\}$ that

$$\begin{aligned} \mathrm{row}(M', 1)_k &= \mathrm{row}(M, i \rightsquigarrow 1(1))_k \\ &= \mathrm{row}(M, i)_k \\ &= \delta_{j,k} \end{aligned}$$

and

$$\begin{aligned} \mathrm{col}(M', j)_k &= \mathrm{row}(M', k)_j \\ &= \mathrm{row}(M, i \rightsquigarrow_n 1(k))_j \\ &= \begin{cases} \mathrm{row}(M, i)_j \text{ if } k = 1 \\ \mathrm{row}(M, k-1)_j \text{ if } 1 < k \leqslant i \\ \mathrm{row}(M, k)_j \text{ if } i < k \end{cases} \\ &= \begin{cases} \delta_{j,j} \text{ if } k = 1 \\ \mathrm{col}(M, j)_{k-1} \text{ if } 1 < k \leqslant i \\ \mathrm{col}(M, j)_k \text{ if } i < k \end{cases} \\ &= \begin{cases} 1 \text{ if } k = 1 \\ \delta_{i,k-1} \text{ if } 1 < k \leqslant i \Rightarrow k - 1 \neq i \\ \delta_{i,k \text{ if } i < k \Rightarrow k \neq i} \end{cases} \\ &= \begin{cases} 1 \text{ if } k = 1 \\ 0 \text{ if } 1 < k \leqslant i \\ 0 \text{ if } i < k \end{cases} \\ &= \delta_{1,k} \end{aligned}$$

By using 10.331 we have then that

$$\det(M') = (-1)^{j+1} \det([1 \boxplus j]M') \tag{10.102}$$

Now if $k, l \in \{1, \ldots, n-1\}$ then we have for $([1 \boxplus j]M')_{k,l}$ the following cases to consider for k, l :

a. $\boldsymbol{(k < i \wedge l < j)}$ then

$$\begin{array}{rcl} ([1 \boxplus j]M')_{k,l} & \underset{1 \leqslant k \wedge l < j}{=} & M'_{k+1,l} \\ & = & \text{row}(M', k+1)_l \\ & = & \text{row}(M, i \rightsquigarrow_n 1(k+1))_l \\ & \underset{1 < k+1 \leqslant i}{=} & \text{row}(M, k)_l \\ & = & M_{k,l} \\ & \underset{k < i \wedge l < j}{=} & ([i \boxplus j]M)_{k,l} \end{array}$$

b. $\boldsymbol{(k < i \wedge j \leqslant l \leqslant n-1)}$ then

$$\begin{array}{rcl} ([1 \boxplus j]M')_{k,l} & \underset{1 \leqslant k \wedge j \leqslant l \leqslant n-1}{=} & M'_{k+1,l+1} \\ & = & \text{row}(M', k+1)_{l+1} \\ & = & \text{row}(M, i \rightsquigarrow_n 1(k+1))_{l+1} \\ & \underset{1 < k+1 \leqslant i}{=} & \text{row}(M, k)_{l+1} \\ & = & M_{k,l+1} \\ & \underset{k < i \wedge j \leqslant l \leqslant n-1}{=} & ([i \boxplus j]M)_{k,l} \end{array}$$

c. $\boldsymbol{(i \leqslant k \leqslant n-1 \wedge l < j)}$ then

$$\begin{array}{rcl} ([1 \boxplus j]M')_{k,l} & \underset{1 \leqslant k \wedge l < j}{=} & M'_{k+1,l} \\ & = & \text{row}(M', k+1)_l \\ & = & \text{row}(M, i \rightsquigarrow_n 1(k+1))_l \\ & \underset{i < k+1}{=} & \text{row}(M, k+1)_l \\ & = & M_{k+1,l} \\ & \underset{i \leqslant k \leqslant n-1 \wedge l < j}{=} & ([i \boxplus j]M)_{k,l} \end{array}$$

d. $\boldsymbol{(i \leqslant k \leqslant n-1 \wedge j \leqslant l \leqslant n-1)}$ then

$$\begin{array}{rcl} ([1 \boxplus j]M')_{k,l} & \underset{1 \leqslant k \wedge j \leqslant l \leqslant n-1}{=} & M'_{k+1,l+1} \\ & = & \text{row}(M', k+1)_{l+1} \\ & = & \text{row}(M, i \rightsquigarrow_n 1(k+1))_{l+1} \\ & \underset{i < k+1}{=} & \text{row}(M, k+1)_{l+1} \\ & = & M_{k+1,l+1} \\ & \underset{i \leqslant k \leqslant n-1 \wedge j \leqslant l \leqslant n-1}{=} & ([i \boxplus j]M)_{k,l} \end{array}$$

So we have that $[1 \boxplus j]M' = [i \boxplus j]M$ and thus by 10.102 we have

$$\det(M') = (-1)^{j+1} \det([i \boxplus j]M) \tag{10.103}$$

Now if Δ is a determinant function in F^n such that $\Delta(\varepsilon_1, \ldots, \varepsilon_n) = 1$ we have by 10.321 that

$$\begin{array}{rcl}\det(M') & = & \Delta(\mathrm{row}(M', 1), \ldots, \mathrm{row}(M', n)) \\ & = & \Delta(\mathrm{row}(M, i \rightsquigarrow_n 1(1)), \ldots, \mathrm{row}(M, i \rightsquigarrow_n 1(n))) \\ & \underset{\text{skew symmetry and } 10.96}{=} & (-1)^{i-1} \cdot \Delta(\mathrm{row}(M, 1), \ldots, \mathrm{row}(M, n)) \\ & \underset{10.321}{=} & (-1)^{i-1} \cdot \det(M)\end{array}$$

Using the fact that $\frac{1}{(-1)^{-1}} = (-1)^{i-1} = (-1)^2 \cdot (-1)^{i-1} = (-1)^{i+1}$ we have $\det(M) = (-1)^{i+1}\det(M')$ and this together with 10.103 gives finally $\det(M) = (-1)^{i+1}(-1)^{j+1}\det([i \boxplus j]M)$ or $\det(M) = (-1)^{i+j}\det([i \boxplus j]M)$ what he have to prove. $\square$

Definition 10.333. *Let F be a field , $n \in \mathbb{N} \setminus \{1\}$ and $M \in M(n \times n, F)$ then we define $[<m]M \in M(m-1 \times m-1, F)$ and $[>m]M \in M(n-m \times n-m, F)$ by*

$$1 < m \leqslant n \Rightarrow ([<m]M)_{k,l} = M_{k,l} \text{ where } k, l \in \{1, \ldots, m-1\}$$
$$1 \leqslant m < n \Rightarrow ([>m]M)_{k,l} = M_{k+m,l+m} \text{ where } k, l \in \{1, \ldots, n-m\}$$

Lemma 10.334. *Let F be a field, $n \in \mathbb{N} \setminus \{1\}$ then we have for $M \in M(n \times n, F)$ then the following holds:*

1. $[1 \boxplus 1]M = [>1]M$
2. $[n \boxplus n]M = [<n]M$
3. *If* $1 < m_1 \leqslant n$ *and* $1 < m_2 \leqslant m_1 - 1$ *then* $[<m_2]([<m_1]M) = [<m_2]M$
4. *If* $1 \leqslant m_1 < n$ *and* $1 \leqslant m_2 < n - m_1$ *then* $[>m_2]([>m_1]M) = [>m_1 + m_2]M$
5. *If* $1 < m \leqslant n$ *then* $([<m]M)^T = [<m]M^T$
6. *If* $1 \leqslant m < n$ *then* $([>m]M)^T = [>m]M^T$

Proof.

1. $\forall i, j \in \{1, \ldots, n-1\}$ we have

$$\begin{array}{rcl}([1 \boxplus 1]M)_{i,j} & \underset{1 \leqslant i,j}{=} & M_{i+1,j+1} \\ & = & ([>1]M)_{i,j}\end{array}$$

so we have $[1 \boxplus 1]M = [>1]M$

2. $\forall i, j \in \{1, \ldots, n-1\}$ we have

$$\begin{array}{rcl}([n \boxplus n]M)_{i,j} & \underset{i,j \leqslant n-1 < n}{=} & M_{i,j} \\ & \underset{1 \leqslant i,j \leqslant n-1}{=} & ([<n]M)_{i,j}\end{array}$$

so we have $[n \boxplus n]M = [<n]M$

3. If $1 \leqslant i, j \leqslant m_2 - 1$ then we have

$$\begin{array}{rcl}([<m_2]([<m_1]M))_{i,j} & \underset{1 \leqslant i,j \leqslant m_2-1}{=} & ([<m_1]M)_{i,j} \\ & \underset{1 \leqslant i,j \leqslant m_2-1 < m_2 \leqslant m_1-1}{=} & M_{i,j} \\ & \underset{1 \leqslant i,j \leqslant m_2-1}{=} & ([<m_2]M)_{i,j}\end{array}$$

so we have $[<m_2]([<m_1]M) = [<m_2]M$

4. First note that $[>m_1]M \in M(n-m_1, n-m_1, F) \Rightarrow [>m_2]([>m_1]M) \in M(n-m_1-m_2 \times n-m_1-m_2, F) = M(n-(m_1+m_2) \times n-(m_1+m_2), F) \ni [>m_1+m_2]M$, so the dimensions are ok. Next if $1 \leqslant i, j \leqslant n-(m_1+m_2)$ then we have

$$\begin{aligned} ([>m_2]([>m_1]M))_{i,j} &= ([>m_1]M)_{i+m_2,j+m_2} \\ &= M_{i+m_2+m_1,j+m_2+m_1} \\ &= M_{i+(m_1+m_2),j+(m_1+m_2)} \\ &= ([>m_1+m_2]M)_{i,j} \end{aligned}$$

so $[>m_2]([>m_1]M) = [>m_1+m_2]M$

5. For $1 \leqslant i, j \leqslant m-1$ we have

$$\begin{aligned} ([<m]M)^T_{i,j} &= ([<m]M)_{j,i} \\ &= M_{j,i} \\ &= M^T_{i,j} \\ &= ([<m]M^T)_{i,j} \end{aligned}$$

so $([<m]M)^T = [<m]M^T$

6. For $1 \leqslant i, j \leqslant n-m$ we have

$$\begin{aligned} ([>m]M)^T_{i,j} &= ([>m]M)_{j,i} \\ &= M_{j+m,i+m} \\ &= M^T_{i+m,j+m} \\ &= ([>m]M^T)_{i,j} \end{aligned}$$

so $([>m]M)^T = [>m]M^T$ $\square$

Definition 10.335. *Given $n, m \in \mathbb{N}$, F a field and $M \in M(n \times m, F)$ then if $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, m\}$ we define $[i \oplus j]M \in M(n \times m, F)$ by $\forall k \in \{1, \ldots, n\}$, $\forall l \in \{1, \ldots, m\}$ we have $([i \oplus j]M)_{k,l} = \left\{ \begin{smallmatrix} M_{k,l} \text{ if } l \neq i \\ \delta_{k,j} \text{ if } l = i \end{smallmatrix} \right.$. Essentially we replace the i-the column by the column $\begin{pmatrix} \delta_{1,j} \\ \ldots \\ \delta_{n,j} \end{pmatrix}$*

Definition 10.336. *Given $n \in \mathbb{N}$, F a field and $M \in M(n \times n, F)$ then $\text{adjoint}(M) \in M(n \times n, F)$ is defined by $\text{adjoint}(M)_{i,j} = \det([i \oplus j]M)$. $\text{adjoint}(M)$ is called the adjoint of the matrix M and $\text{adjoint}(M)_{i,j} = \det([i \oplus j]M)$ is the i, j-the co-factor of M.*

Example 10.337. If $n = 1$ then $\text{adjoint}(M) = \left(\text{adjoint}(M)_{1,1} \right) = (\det([1 \oplus 1]\left(m_{1,1} \right))) = \left(\det(1) \right) = \left(1 \right)$

Theorem 10.338. *If X is a n-dimensional vector space $(n \in \mathbb{N})$ over a field F of a characterization zero, $L \in \text{Hom}(X, X)$ a linear transformation and $\{e_i\}_{i \in \{1, \ldots, n\}}$ a basis of X then $\mathcal{M}(\text{adjoint}(L), \{e_i\}_{i \in \{1, \ldots, n\}}, \{e_i\}_{i \in \{1, \ldots, n\}}) = \text{adjoint}(\mathcal{M}(L, \{e_i\}_{i \in \{1, \ldots, n\}}, \{e_i\}_{i \in \{1, \ldots, n\}}))$. Essentially the matrix of the adjoint of a linear transformation is the adjoint of the matrix of a linear transformation.*

Proof. Let Δ be a determinant function in X such that $\Delta(e_1, \ldots, e_n) = 1$ then we have

$$\begin{aligned}
\sum_{k \in \{1,\ldots,n\}} \mathcal{M}(\text{adjoint}(L), \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})_{k,j} \cdot e_k &\underset{\text{definition of } \mathcal{M}}{=} \text{adjoint}(L)(e_j) \\
&\underset{\Delta(e_1,\ldots,e_n)=1}{=} \Delta(e_1, \ldots, e_n) \cdot \text{adjoint}(L)(e_j) \\
&\underset{10.285}{=} \breve{\Delta}_L(e_1, \ldots, e_n)(e_j) \\
&= \sum_{k\in\{1,\ldots,n\}} (-1)^{k-1} \cdot \Delta(e_j, L(e_1), \ldots, L(e_{k-1}), L(e_{k+1}), \ldots, L(e_n)) \cdot e_k
\end{aligned}$$

From the uniqueness of a expansion in a basis we have then that

$$\mathcal{M}(\text{adjoint}(L), \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})_{k,j} = (-1)^{k-1}\Delta(e_j, L(e_1), \ldots, L(e_{k-1}), L(e_{k+1}), \ldots, L(e_n)) \tag{10.104}$$

Define now $\{b_i^{k,j}\}_{i\in\{1,\ldots,n\}}$ by $b_i^{j,k} = \begin{cases} L(e_i) \text{ if } i \neq k \\ e_j \text{ ik } i = k \end{cases}$ then we have that

$$\begin{aligned}
\Delta(e_j, L(e_1), \ldots., L(e_{k-1}), L(e_{k+1}), \ldots, L(e_n)) &= \Delta(b_k^{k,j}, b_1^{k,j}, \ldots, b_{k-1}^{k,j}, b_{k+1}^{k,j}, \ldots, b_n^{k,j}) \\
&= \Delta(b_{1\rightsquigarrow_n k(1)}^{k,j}, b_{1\rightsquigarrow_n k(2)}^{k,j}, \ldots, b_{1\rightsquigarrow_n k(k)}^{k,j}, b_{1\rightsquigarrow_n k(k+1)}^{k,j}, \ldots, b_{1\rightsquigarrow_n (n)}^{k,j}) \\
&= (1 \rightsquigarrow_n k)\Delta(b_1^{k,j}, \ldots, b_n^{k,j}) \\
&= (-1)^{k-1}\Delta(b_1^{k,j}, \ldots, b_n^{k,j})
\end{aligned}$$

Using the above and 10.104 we have

$$\mathcal{M}(\text{adjoint}(L), \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})_{k,j} = \Delta(b_1^{k,j}, \ldots, b_n^{k,j}) \tag{10.105}$$

Define now

Define now $\forall i, j \in \{1, \ldots, n\}$ the matrix $M^{i,j} \in M(n \times n, F)$ by $\forall r, s \in \{1, \ldots, n\}$ we have

$$(M^{i,j})_{r,s} = \begin{cases} L(e_s)_r \text{ if } s \neq i \\ \delta_{j,r \text{ if } s=i} \end{cases} \tag{10.106}$$

so that $(M^{k,j})_{r,s} = \begin{cases} (b_s^{k,j})_r \text{ if } s \neq k \\ (e_j)_r \text{ if } s = k \end{cases} = \begin{cases} (b_s^{k,j})_r \text{ if } s \neq k \\ (b_s^{k,j})_r \text{ if } s = k \end{cases} = (b_s^{k,j})_r$. If we take now the linear mapping $L^{i,j}: X \to X$ by $L^{i,j} = \mathcal{M}(\{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})^{-1}(M^{i,j})$ then by 10.307 and the fact that $e_t = \sum_{l\in\{1,\ldots,n\}} \delta_{l,t} \cdot e_l$ we have that

$$\begin{aligned}
L^{i,j}(e_t) &= \sum_{r\in\{1,\ldots,n\}} \left(\sum_{s\in\{1,\ldots,n\}} (M^{i,j})_{r,s} \cdot \delta_{t,s} \right) \cdot e_r \\
&= \sum_{r\in\{1,\ldots,n\}} (M^{i,j})_{r,t} \cdot e_r
\end{aligned}$$

If now $i = k$ then

$$\begin{aligned} L^{k,j}(e_i) &= \sum_{r\in\{1,\ldots,n\}} (M^{i,j})_{r,i}\cdot e_r \\ &= \sum_{r\in\{1,\ldots,n\}} (b_i^{k,j})_r\cdot e_r \\ &= b_i^{k,j} \end{aligned}$$

This together with 10.105 gives

$$\begin{aligned} \mathcal{M}(\text{adjoint}(L), \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})_{k,j} &= \Delta(L^{k,j}(e_1),\ldots,L^{k,j}(e_n)) \\ &= \det(L_{k,j})\cdot\Delta(e_1,\ldots,e_n) \\ &= \det(L_{k,j}) \\ &\underset{10.319}{=} \det(M^{k,j}) \end{aligned}$$

so we have

$$\mathcal{M}(\text{adjoint}(L), \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})_{k,j} = \det(M^{k,j}) \tag{10.107}$$

Now for $r, s\in\{1,\ldots,n\}$ we have the following cases to consider for $(M^{i,j})_{r,s}$

1. $(s \neq i)$ then

$$\begin{aligned} (M^{i,j})_{r,s} &= L(e_s)_r \\ &= \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})_{r,s} \\ &= ([i\oplus j]\mathcal{M}(L, \{e_i\}_{i\in[1,\ldots,n]}, \{e_i\}_{i\in\{1,\ldots,n\}})) \end{aligned}$$

2. $(s = k)$ then

$$\begin{aligned} (M^{i,j})_{r,s} &= \delta_{j,r} \\ &= ([i\oplus j]\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}}))_{r,s} \end{aligned}$$

So that $M^{i,j} = [i\oplus j]\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})$. Using this and 10.107 gives us $\mathcal{M}(\text{adjoint}(L), \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})_{k,j} = \det([k\oplus j]\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})) \underset{\text{definition of adjoint of a matrix}}{=} \text{adjoint}(\mathcal{M}(L.\{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}}))_{k,j}$ or finally

$\mathcal{M}(\text{adjoint}(L), \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}}) = \text{adjoint}(\mathcal{M}(L.\{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}}))$ □

Theorem 10.339. *If X is a n-dimensional vector space ($n\in\mathbb{N}$) over a field F of characterization 0, $L\in\operatorname{Hom}(X,X)$ and $\{e_i\}_{i\in\{1,\ldots,n\}}$ a basis of X then we have the following*

1. $\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})\cdot\text{adjoint}(\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})) = \det(\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}}))\cdot E$
2. $\text{adjoint}(\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}}))\cdot\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}}) = \det(\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}}))\cdot E$

or if we write out the matrix products directly we have

1. $\sum_{k\in\{1,\ldots,n\}} \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})_{i,k} \cdot \mathrm{adjoint}(\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}}))_{k,j} = \det(\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})) \cdot \delta_{i,j}$
2. $\sum_{k\in\{1,\ldots,n\}} \mathrm{adjoint}(\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}}))_{i,k} \cdot \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})_{k,j} = \det(\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})) \cdot \delta_{i,j}$

Proof.

1. *Using* 10.287 *we have that* $L \circ \mathrm{adjoint}(L) = \det(L) \cdot 1_X$ *so that* $\mathcal{M}(L \circ \mathrm{adjoint}(L), \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,\}}) = \mathcal{M}(\det(L) \cdot 1_X, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,\}}) \underset{10.305}{\Rightarrow} \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}}) \cdot \mathcal{M}(\mathrm{adjoint}(L), \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,\}}) = \mathcal{M}(\det(L) \cdot 1_X, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,\}}) \underset{10.304 \text{ and } 10.302}{=} \det(L) \cdot E \underset{10.319 \text{ and } 10.338}{\Rightarrow} \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}}) \cdot \mathrm{adjoint}(\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})) = \det(\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})) \cdot E$
2. *Using* 10.287 *we have that* $\mathrm{adjoint}(L) \circ L = \det(L) \cdot 1_X$ *so that* $\mathcal{M}(\mathrm{adjoint}(L) \circ L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,\}}) = \mathcal{M}(\det(L) \cdot 1_X, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,\}}) \underset{10.305}{\Rightarrow} \mathcal{M}(\mathrm{adjoint}(L), \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}}) \cdot \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,\}}) = \mathcal{M}(\det(L) \cdot 1_X, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,\}}) \underset{10.304 \text{ and } 10.302}{=} \det(L) \cdot E \underset{10.319 \text{ and } 10.338}{\Rightarrow} \mathrm{adjoint}(\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})) \cdot \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}}) = \det(\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})) \cdot E$ □

The above theorem gives the reformulation of 10.287 in matrix form, the next theorem works just for matrices alone (which is made possible by the bijection between $M(m \times n, F)$ and $\mathrm{Hom}(F^n, F^m)$ (see 10.307 and 10.313).

Theorem 10.340. *If F is a field of characterization zero then if $n \in \mathbb{N}$ we have for $M \in M(n \times n, F)$ that*

1. $M \cdot \mathrm{adjoint}(M) = \det(M) \cdot E$ *(or* $\sum_{k\in\{1,\ldots,n\}} M_{i,k} \cdot \mathrm{adjoint}(M)_{k,j} = \det(M) \cdot \delta_{i,j}$
2. $\mathrm{adjoint}(M) \cdot M = \det(M) \cdot E$ *(or* $\sum_{k\in\{1,\ldots,n\}} \mathrm{adjoint}(M)_{i,k} \cdot M_{k,j} = \det(M) \cdot \delta_{i,j}$
3. *If $n > 1$ then* $\mathrm{adjoint}(M)_{i,j} = (-1)^{i+j} \cdot \det([j \boxplus i]M)$
4. *If $n > 1$ then* $\det(M) = \sum_{i\in\{1,\ldots n\}} (-1)^{i+j} M_{i,j} \cdot \det([i \boxplus j]M)$ *where* $j \in \{1, \ldots, n\}$ *(this is the expansion of the determinant with respect to the j-the column).*
5. *If $n > 1$ then* $\det(M) = \sum_{j\in\{1,\ldots,n\}} (-1)^{i+j} M_{i,j} \cdot \det([i \boxplus j]M)$ *where* $i \in \{1, \ldots, n\}$ *(this is the expansion of the determinant with respect to the i-the row)*
6. $\det(E) = 1$

7. *If* $M \in M(1 \times 1, F)$ *then* $\operatorname{adjoint}(M) = (1)$ *and* $\det(M) = M_{1,1}$

Proof.

1. Let $\mathcal{M}(\{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}})\colon \operatorname{Hom}(F^n, F^n) \to M(n \times n, F)$ be the bijection defined in 10.307. If $M \in M(n \times n, F)$ take then $L = \mathcal{M}(\{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}})^{-1}(M)$ then we have by 10.339 that $\mathcal{M}(L, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}}) \cdot \operatorname{adjoint}(\mathcal{M}(L, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}})) = \det(\mathcal{M}(L, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}})) \cdot E$ which as $M = \mathcal{M}(\{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}})(L) = \mathcal{M}(L, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}})$ means that $M \cdot \operatorname{adjoint}(M) = \det(M) \cdot E$.

2. Let $\mathcal{M}(\{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}})\colon \operatorname{Hom}(F^n, F^n) \to M(n \times n, F)$ be the bijection defined in 10.307. If $M \in M(n \times n, F)$ take then $L = \mathcal{M}(\{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}})^{-1}(M)$ then we have by 10.339 that $\operatorname{adjoint}(\mathcal{M}(L, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}})) \cdot \mathcal{M}(L, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}}) = \det(\mathcal{M}(L, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}})) \cdot E$ which as $M = \mathcal{M}(\{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}})(L) = \mathcal{M}(L, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}}, \{\varepsilon_i\}_{i\in\{1,\ldots,n\}})$ means that $\operatorname{adjoint}(M) \cdot M = \det(M) \cdot E$.

3. Note first that if $r, s \in \{1, \ldots, n\}$ then we have

$$\begin{aligned}
\det([r \oplus s]M) &= \det(([r \oplus s]M)^T) \\
&= \det(([r \oplus s]M)^T) \cdot \delta_{r,r} \\
&\underset{\text{see (1)}}{=} \sum_{k\in\{1,\ldots,n\}} ([r \oplus s]M)^T_{r,k} \cdot \operatorname{adjoint}(([r \oplus s]M)^T)_{k,r} \\
&= \sum_{k\in\{1,\ldots,n\}} ([r \oplus s]M)_{k,r} \cdot \operatorname{adjoint}(([r \oplus s]M)^T)_{k,r} \\
&\underset{([r\oplus s]M)_{k,r} = \delta_{k,s}}{=} \sum_{k\in\{1,\ldots,n\}} \delta_{k,s} \cdot \operatorname{adjoint}(([r \oplus s]M)^T)_{k,r} \\
&= \operatorname{adjoint}(([r \oplus s]M)^T)_{s,r} \\
&\underset{\text{definition of adjoint}}{=} \det([s \oplus r]([r \oplus s]M)^T)
\end{aligned}$$

this gives

$$\det([r \oplus s]M) = \det([s \oplus r]([r \oplus s]M)^T) \tag{10.108}$$

If now $j \in \{1, \ldots, n\}$ then we have

$$\begin{aligned}
\operatorname{row}([s \oplus r]([r \oplus s]M)^T, r)_j &= ([s \oplus r]([r \oplus s]M)^T)_{r,j} \\
&= \begin{cases} \delta_{r,r} = 1 \text{ if } j = s \\ ([r \oplus s]M)^T_{r,j} \text{ if } j \neq s \end{cases} \\
&= \begin{cases} 1 \text{ if } j = s \\ ([r \oplus s]M)_{j,r} \text{ if } j \neq s \end{cases}
\end{aligned}$$

$$
\begin{aligned}
&= \begin{cases} 1 \text{ if } j = s \\ \delta_{j,s} \text{ if } j \neq s \end{cases} \\
&= \begin{cases} 1 \text{ if } j = s \\ 0 \text{ if } j \neq s \end{cases} \\
&= \delta_{j,s}
\end{aligned}
$$

giving

$$\mathrm{row}([s \oplus r]([r \oplus s]M)^T, r)_j = \delta_{j,s} \tag{10.109}$$

Also we have that

$$
\begin{aligned}
\mathrm{col}([s \oplus r]([r \oplus s]M)^T, s)_j &= ([s \oplus r]([r \oplus s]M)^T)_{j,s} \\
&= \delta_{j,r}
\end{aligned}
$$

giving

$$\mathrm{col}([s \oplus r]([r \oplus s]M)^T, s)_j = \delta_{j,r} \tag{10.110}$$

Using 10.109, 10.110 and 10.332 we have then

$$\det([s \oplus r]([r \oplus s]M)^T) = (-1)^{r+s}\det([r \boxplus s]([s \oplus r]([r \oplus s]M)^T)) \tag{10.111}$$

Consider now the following cases for $i, j \in \{1, \ldots, n-1\}$ in $([r \boxplus s]([s \oplus r]([r \oplus s]M)^T))_{i,j}$:

a. $\boldsymbol{(i < r \wedge j < s)}$ then

$$
\begin{aligned}
([r \boxplus s]([s \oplus r]([r \oplus s]M)^T))_{i,j} &= (([s \oplus r]([r \oplus s]M)^T))_{i,j} \\
&\underset{j<s \Rightarrow j \neq s}{=} ([r \oplus s]M)^T_{i,j} \\
&= ([r \oplus s]M)_{j,i} \\
&\underset{j<s \Rightarrow j \neq s}{=} M_{j,i} \\
&= M^T_{i,j} \\
&\underset{i<r \wedge j<s}{=} ([r \boxplus s]M^T)_{i,j}
\end{aligned}
$$

b. $\boldsymbol{(r \leqslant i \leqslant n-1 \wedge j < s)}$ then

$$
\begin{aligned}
([r \boxplus s]([s \oplus r]([r \oplus s]M)^T))_{i,j} &= (([s \oplus r]([r \oplus s]M)^T))_{i+1,j} \\
&\underset{j<s \Rightarrow j \neq s}{=} ([r \oplus s]M)^T_{i+1,j} \\
&= ([r \boxplus s]M)_{j,i+1} \\
&\underset{r \leqslant i \Rightarrow r \neq i+1}{=} M_{j,i+1} \\
&= M^T_{i+1,j} \\
&\underset{r \leqslant i < n-1, j<s}{=} ([r \boxplus s]M^T)_{i,j}
\end{aligned}
$$

c. $\boldsymbol{(i<r\wedge s\leqslant j\leqslant n-1)}$ then

$$\begin{array}{rcl}([r\boxplus s]([s\oplus r]([r\oplus s]M)^T))_{i,j} & = & (([s\oplus r]([r\oplus s]M)^T))_{i,j+1}\\ & \underset{s\leqslant j\Rightarrow s\neq j+1}{=} & ([r\oplus s]M)^T_{i,j+1}\\ & = & ([r\boxplus s]M)_{j+1,i}\\ & \underset{i<r\Rightarrow i\neq r}{=} & M_{j+1,i}\\ & = & M^T_{i,j+1}\\ & \underset{i<r\wedge s\leqslant j\leqslant n-1}{=} & ([r\boxplus s]M^T)_{i,j}\end{array}$$

d. $\boldsymbol{(r\leqslant i\leqslant n-1, s\leqslant j\leqslant n-1)}$ then

$$\begin{array}{rcl}([r\boxplus s]([s\oplus r]([r\oplus s]M)^T))_{i,j} & = & (([s\oplus r]([r\oplus s]M)^T))_{i+1,j+1}\\ & \underset{s\leqslant j\Rightarrow s\neq j+1}{=} & ([r\oplus s]M)^T_{i+1,j+1}\\ & = & ([r\oplus s]M)_{j+1,i+1}\\ & \underset{r\leqslant i\Rightarrow r\neq i+1}{=} & M_{j+1,i+1}\\ & = & M^T_{i+1,j+1}\\ & = & ([r\boxplus s]M^T)_{i,j}\end{array}$$

Which means that $([r\boxplus s]([s\oplus r]([r\oplus s]M)^T))=([r\boxplus s]M^T)\underset{10.329}{=}([s\boxplus r]M)^T$, This together with $\operatorname{adjoint}(M)_{r,s}=\det([r\oplus s]M^T)$, 10.108 and 10.111 gives that

$$\operatorname{adjoint}(M)_{r,s}=(-1)^{r+s}\det(([s\boxplus r]M)^T)\underset{10.318}{=}(-1)^{r+s}\det([s\boxplus r]M)$$

4. If $i\in\{1,\ldots,n\}$ then we have

$$\begin{array}{rcl}\det(M) & = & \det(M)\cdot\delta_{i,i}\\ & \underset{\text{see (2)}}{=} & \sum_{k\in\{1,\ldots,n\}}\operatorname{adjoint}(M)_{i,k}\cdot M_{k,i}\\ & \underset{\text{see (3)}}{=} & \sum_{k\in\{1,\ldots,n\}}(-1)^{i+k}\det([k\boxplus i]M)\cdot M_{k,i}\\ & = & \sum_{k\in\{1,\ldots,n\}}(-1)^{i+k}M_{k,i}\cdot\det([k\boxplus i]M)\end{array}$$

5. If $i\in\{1,\ldots,n\}$ then we have

$$\begin{array}{rcl}\det(M) & = & \det(M)\cdot\delta_{i,i}\\ & \underset{\text{see (1)}}{=} & \sum_{k\in\{1,\ldots,n\}}M_{i,k}\cdot\operatorname{adjoint}(M)_{k,i}\\ & \underset{\text{see (3)}}{=} & \sum_{k\in\{1,\ldots,n\}}(-1)^{i+k}M_{i,k}\cdot\det([i\boxplus k]M)\end{array}$$

6. This follows from 10.320.

7. This follows from 10.337 and 10.315 □

Using the above theorem we can calculate the determinant and adjoint of every matrix.

Example 10.341. Using (4) from the above theorem and $i=1$ then

$$\begin{aligned}\begin{vmatrix} -2 & 2 & 3 \\ -1 & 1 & 3 \\ 2 & 0 & 1 \end{vmatrix} &= (-1)^{1+1} M_{1,1} \det([1 \boxplus 1]M) + (-1)^{1+2} M_{2,1} \det([2 \boxplus 1]M) + (-1)^{1+3} M_{3,1} \det([3 \boxplus 1]M) \\ &= -2 \cdot \begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix} + (-1) \cdot (-1) \cdot \begin{vmatrix} 2 & 3 \\ 0 & 1 \end{vmatrix} + 2 \cdot \begin{vmatrix} 2 & 3 \\ 1 & 3 \end{vmatrix} \\ &= -2 \cdot (1 \cdot |\ 1\ | + (-1) \cdot 0 \cdot |\ 3\ |) + (2 \cdot |\ 1\ | + (-1) \cdot 0 \cdot |\ 3\ |) + 2 \cdot (2 \cdot |\ 3\ | + (-1) \cdot 1 \cdot |\ 3\ |) \\ &= -2 + 2 + 2 \cdot 3 = 6\end{aligned}$$

Example 10.342. Using (3) from the above we have

$$\begin{aligned}\operatorname{adjoint}\left(\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}\right) &= \begin{pmatrix} 1 \cdot \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} & -1 \cdot \begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} & 1 \cdot \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} \\ -1 \cdot \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} & 1 \cdot \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} & -1 \cdot \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} \\ 1 \cdot \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} & -1 \cdot \begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix} & 1 \cdot \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} \end{pmatrix} \\ &= \begin{pmatrix} -3 & 6 & -3 \\ 6 & -12 & 6 \\ -3 & 6 & -3 \end{pmatrix}\end{aligned}$$

The following corollary allows us to calculate in a much faster way the determinant for matrices where some of the leading (trailing) rows (columns) are unit rows (columns).

Corollary 10.343. *If F is a field of characterization zero then if $n \in \mathbb{N}$ we have the following for $M \in M(n \times n, F)$*

1. *If $\exists m \in \mathbb{N} \vdash 1 \leqslant m < n$ such that $\forall i \vdash i \leqslant m$ we have $\operatorname{col}(M, i)_j = \delta_{i,j}$ then* $\det(M) = \det([>m]M)$

2. *If $\exists m \in \mathbb{N} \vdash 1 < m \leqslant n$ such that $\forall i \vdash m \leqslant i \leqslant n$ we have $\operatorname{col}(M, i)_j = \delta_{i,j}$ then* $\det(M) = \det([<m]M)$

3. *If $\exists m \in \mathbb{N} \vdash 1 \leqslant m < n$ such that $\forall i \vdash i \leqslant m$ we have $\operatorname{row}(M, i)_j = \delta_{i,j}$ then* $\det(M) = \det([>m]M)$

4. *If $\exists m \in \mathbb{N} \vdash 1 < m \leqslant n$ such that $\forall i \vdash m \leqslant i \leqslant n$ we have $\operatorname{row}(M, i)_j = \delta_{i,j}$ then* $\det(M) = \det([<m]M)$

Proof.

1. We use induction on m to prove this, so let $E_n = \{m \in \mathbb{N} | \text{if } 1 \leqslant m < n \text{ and } \forall i \in \mathbb{N} \vdash i \leqslant m \text{ we have } \text{col}(M, i)_j = \delta_{i,j} \text{ then } \det(M) = \det([>m]M)\}$ then we have:

   a. If $m = 1$ then by the previous theorem we have

$$\begin{aligned}
\det(M) &= \sum_{i \in \{1,\ldots,n\}} (-1)^{i+1} M_{i,1} \cdot \det([i \boxplus 1]M) \\
&= \sum_{i \in \{1,\ldots,n\}} (-1)^{i+1} \text{col}(M,1)_i \cdot \det([i \boxplus 1]M) \\
&= \sum_{i \in \{1,\ldots,n\}} (-1)^{i+1} \cdot \delta_{1,i} \det([i \boxplus 1]M) \\
&= (-1)^{1+1} \det([1 \boxplus 1]M) = \det([1 \boxplus 1]M) \\
&= \det([>1]M)
\end{aligned}$$

   so we have $1 \in E_n$

   b. If $m \in E_n$ then if $m+1$ is such that $1 \leqslant m+1 < n$ and $\forall i \vdash i \leqslant m+1$ we have $\text{col}(M, i)_j = \delta_{i,j}$. We have then as $m \in E_n \subseteq \mathbb{N} \Rightarrow 1 \leqslant m < m+1 < n$ and $\forall i \vdash i \leqslant m$ we have $\text{col}(M, i) = \delta_{i,j}$ so as $m \in E_n$ we have

$$\det(M) = \det([>m]M). \tag{10.112}$$

   Now if $j \in \{1, \ldots, n-m\}$ then $\text{col}([>m]M, 1)_j = ([>m]M)_{j,1} = M_{j+m,1+m} = \text{col}(M, 1+m)_{j+m} = \delta_{m+1,j+m}$. Using the previous theorem we have then

$$\begin{aligned}
\det([>m]M) &= \sum_{i \in \{1,\ldots,n-m\}} (-1)^{i+1} ([>m]M)_{i,1} \cdot \\
&\quad \det([i \boxplus 1]([>m]M)) \\
&= \sum_{i \in \{1,\ldots,n-m\}} (-1)^{i+1} \text{col}([>m]M, 1)_i \cdot \\
&\quad \det([i \boxplus 1]([>m]M)) \\
&= \sum_{i \in \{1,\ldots,n-m\}} (-1)^{i+1} \delta_{m+1,i+m} \cdot \\
&\quad \det([i \boxplus 1]([>m]M)) \\
&\underset{m+1=i+m \Rightarrow i=1}{=} (-1)^2 \cdot \det([1 \boxplus 1]([>m]M)) \\
&\underset{10.334}{=} \det([>1]([>m]M)) \\
&\underset{10.334}{=} \det([>m+1]M)
\end{aligned}$$

   Using the above with 10.112 gives $\det(M) = \det([>m+1]M)$ and thus $m+1 \in E_n$.

   Using mathematical induction we have then $E_n = \mathbb{N}$ proving (1).

2. We use induction on n to prove this, so let $E=\{n\in\{2,\ldots\}|\text{if } 1<m\leqslant n \text{ and } \forall i\in\mathbb{N}\vdash m\leqslant i\leqslant n \text{ and for } M\in M(n\times n,F) \text{ we have } \operatorname{col}(M,i)_j=\delta_{i,j} \text{ then } \det(M)=\det([<m]M)\}$ then we have:

   a. If $n=2$ then $m=2$ and by the previous theorem we have

$$\begin{array}{rcl}\det(M) & = & \sum_{i\in\{1,\ldots,2\}}(-i)^{i+2}\cdot M_{i,2}\cdot\det([i\boxplus 2]M)\\ & = & \sum_{i\in\{1,\ldots,2\}}(-i)^{i+2}\cdot\operatorname{col}(M,2)_i\cdot\det([i\boxplus 2]M)\\ & = & \sum_{i\in\{1,\ldots,2\}}(-i)^{i+2}\cdot\delta_{i,2}\cdot\det([i\boxplus 2]M)\\ & = & (-1)^{2+2}\cdot\det([2\boxplus 2]M)\\ & = & \det([2\boxplus 2]M)\\ & \underset{10.334}{=} & \det([<2]M)\underset{m=2}{=}\det([<m]M)\end{array}$$

   so we have $2\in E$

   b. Let $n\in E$ then if $1<m\leqslant n+1$ is such that $\forall i\in\mathbb{N}\vdash m\leqslant i\leqslant n+1$ we have $\operatorname{col}(M,i)_j=\delta_{i,j}$ then we have for $[n+1\boxplus n+1]M\underset{10.334}{=}[<m+1]M$ that $\forall i\in\mathbb{N}\vdash m\leqslant i\leqslant n$

$$\begin{array}{rcl}\operatorname{col}([n+1\boxplus n+1]M,i)_j & \underset{i\leqslant n<n+1}{=} & M_{j,i}\\ & = & \operatorname{col}(M,j)_i\\ & \underset{m\leqslant i\leqslant n\Rightarrow m\leqslant i\leqslant n+1}{=} & \delta_{i,j}\end{array}$$

   so as $n\in E$ we have $\det([<m+1]M)\underset{10.334}{=}\det([<m]([<m+1]M))\underset{m\leqslant((m+1)-1)\wedge 10.334}{=}\det([<m])$ so that $n+1\in E$.

   Using mathematical induction we have then $E=\{2,\ldots\}=\mathbb{N}\setminus\{1\}$ proving (2)

3. As $\operatorname{row}(M,i)_j=M_{i,j}=M^T_{j,i}=\operatorname{col}(M^T,i)_j$ we have $\forall i\in\mathbb{N}\vdash i\leqslant m$ that $\operatorname{col}(M^T,i)_j=\delta_{i,j}$ and thus by (1) we have $\det([>m]M^T)\underset{10.334}{=}\det(([>m]M)^T)\underset{10.318}{=}\det([>m]M)$

4. As $\operatorname{row}(M,i)_j=M_{i,j}=M^T_{j,i}=\operatorname{col}(M^T,i)_j$ we have $\forall i\in\mathbb{N}\vdash m\leqslant i\leqslant n$ that $\operatorname{col}(M^T,i)_j=\delta_{i,j}$ and thus by (1) we have $\det([<m]M^T)\underset{10.334}{=}\det(([<m]M)^T)\underset{10.318}{=}\det([<m]M)$ □

Theorem 10.344. *If F is a field of characterization zero then for $n\in\mathbb{N}$ and $M\in M(n\times n,F)$ we have*

$$\begin{array}{c}\det(M)\neq 0\\ \Leftrightarrow\\ \exists M^{-1}\in M(n\times n,F)\text{ with }M^{-1}\cdot M=E=M\cdot M^{-1}\quad\text{where }M^{-1}=\frac{1}{\det(M)}\cdot\operatorname{adjoint}(M)\end{array}$$

Proof. Using 10.340 we have $M\cdot\text{adjoint}(M)=\text{adjoint}(M)\cdot M=\det(M)\cdot E$ and thus we have:

1. **($\det(M)\neq 0$)** then $\det(M)^{-1}=\frac{1}{\det(M)}$ exists and we have $M\cdot\left(\frac{1}{\det(M)}\cdot\text{adjoint}(M)\right)=\left(\frac{1}{\det(M)}\cdot\text{adjoint}(M)\right)\cdot M=\frac{1}{\det(M)}\cdot\det(M)\cdot E=E$. So M^{-1} exists and $M^{-1}=\frac{1}{\det(M)}\cdot\text{adjoint}(M)$
2. **(Assume M^{-1} exists)** then $M\cdot M^{-1}=E$ and $1=\det(E)=\det(M\cdot M^{-1})=\det(M)\cdot\det(M^{-1})\Rightarrow\det(M)\neq 0$ □

10.10 Expansion of linear transformation in elementair transformations

Definition 10.345. *Let $n\in\mathbb{N}$, X be a n-dimensional vector space over a field F of characterization zero with a basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ then a elementair transformation is a **elementair** transformation on X if it is of one of the following forms*

1. $\forall i\in\{1,\ldots,n\}$ define $\sigma_n(i,\alpha)\colon X\to X$ *is defined by* $\forall k\in\{1,\ldots,n\}$ *we have* $\sigma_n(i,\alpha)(e_k)=\begin{cases} e_k \text{ if } k\neq i\\ \alpha\cdot e_i \text{ if } k=i\end{cases}$ *where* $i\in\{1,\ldots,n\}$ *and* $\alpha\in F$
2. $\forall i,j\in\{1,\ldots,n\}$ define $\tau_n(i,j)\colon X\to X$ *is defined by* $\forall k\in\{1,\ldots,n\}$ *we have* $\tau_n(i,j)(e_k)=\begin{cases} e_k \text{ if } k\neq i,j\\ e_j \text{ if } k=i\\ e_i \text{ if } k=j\end{cases}$ *where* $i,j\in\{1,\ldots,n\}$
3. $\forall i,j\in\{1,\ldots,n\}\vdash i\neq j$ define $\beta_n(i,j,\alpha)\colon X\to X$ *is defined by* $\forall k\in\{1,\ldots,n\}$ *we have* $\beta_n(i,j,\alpha)(e_k)=\begin{cases} e_k \text{ if } k\neq i\\ e_i+\alpha\cdot e_j \text{ if } k=i\end{cases}$ where $i,j\in\{1,\ldots,n\}$ *and* $\alpha\in F$

Note 10.346. Let $n\in\mathbb{N}$, X a n-dimensional vector space over a field F of characterization zero with basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ then we have

1. $\forall i\in\{1,\ldots,n\}$, $\forall\alpha\in F$ we have $\det(\sigma_n(i,\alpha))=\alpha$, so $\sigma_n(i,\alpha)$ is a isomorphism if and only if $\alpha\neq 0$
2. $\forall i,j\in\{1,\ldots,n\}$ we have $\det(\tau_n(i,j))=-1$ if $i\neq j$ and $\det(\tau_n(i,j))=1$ if $i=j$, so τ_n is a isomorphism
3. $\forall i,j\in\{1,\ldots,n\}$, $\forall\alpha\in F$ we have $\det(\beta_n(i,j,\alpha))=1$
4. If $i=j$ then $\tau_n(i,j)=1_X$
5. $\tau_n(i,j)\circ\tau_n(i,j)=1_X$ proving that $\tau_n(i,j)=\tau_n(i,j)^{-1}$

Proof.

1. If $i,j\in\{1,\ldots,n\}$ with $i=j$ then we have if $k\in\{1,\ldots,n\}$ that $\tau_n(i,j)(e_k)=\begin{cases} e_k \text{ if } k\in\{1,\ldots,n\}\setminus\{i\}\\ e_j \text{ if } k=i\\ e_i \text{ if } k=i\end{cases}\underset{i=j\Rightarrow e_i=e_j}{=}\begin{cases} e_k \text{ if } k\in\{1,\ldots,n\}\setminus\{i,j\}\\ e_i \text{ if } k=i\\ e_j \text{ if } k=j\end{cases}=e_k=1_X(e_k)$ proving that $\tau_n(i,j)=1_X$
2. Let $i,j\in\{1,\ldots,n\}$ then we have for $k\in\{1,\ldots,n\}$ either

   $\boldsymbol{k\in\{1,\ldots,n\}\setminus\{i,j\}}$**.** then $(\tau_n(i,j)\circ\tau_n(i,j))(e_k)=\tau_n(i,j)(\tau_n(i,j)(e_k))=\tau_n(i,j)(e_k)=e_k=1_X(e_k)$

$k=i$. $(\tau_n(i,j)\circ\tau_n(i,j))(e_k)=\tau_n(i,j)(\tau_n(i,j)(e_k))=\tau_n(i,j)(e_j)=e_i=e_k=1_X(e_k)$

$k=j$. $(\tau_n(i,j)\circ\tau_n(i,j))(e_k)=\tau_n(i,j)(\tau_n(i,j)(e_k))=\tau_n(i,j)(e_i)=e_j=e_k=1_X(e_k)$

which proves that $\tau_n(i,j)\circ\tau_n(i,j)=1_X$ □

Theorem 10.347. *Let $n\in\mathbb{N}$, X be a n-dimensional vectorspace over a fiield F of characterization zero with basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ then we have*

1. *$\forall i\in\{1,\ldots,n\}$, $\forall\alpha\in F$ we have $\det(\sigma_n(i,\alpha))=\alpha$, so $\sigma_n(i,\alpha)$ is a isomorphism if and only if $\alpha\neq 0$*
2. *$\forall i,j\in\{1,\ldots,n\}$ we have $\det(\tau_n(i,j))=-1$ if $i\neq j$ and $\det(\tau_n(i,j))=1$ if $i=j$, so τ_n is a isomorphism*
3. *$\forall i,j\in\{1,\ldots,n\}$, $\forall\alpha\in F$ we have $\det(\beta_n(i,j,\alpha))=1$*

Proof. See the following examples 10.277. 10.278 and 10.279, together with the fact that if $i=j$ we have that $\tau_n(i,j)=1_X$. □

The following theorem show how to extend a elementair function.

Definition 10.348. *Let $n\in\mathbb{N}\setminus\{1\}$, X a n-dimensional vectorspace over a field F of characterization zero with basis $\{e_i\}_{i\in\{1,\ldots,n\}}$, $Y=S(\{e_i|i\in\{1,\ldots,n-1\}\})$ (see 10.128) a $(n-1)$-dimensional subvector space of X then if $L\in\mathrm{Hom}(Y,Y)$ we define $L_{\langle n\rangle}\in\mathrm{Hom}(X,X)$ by $L_{\langle n\rangle}(e_k)=\left\{\begin{smallmatrix} L(e_k) \text{ if } k\in\{1,\ldots,n-1\}\\ e_n \text{ if } k=n\end{smallmatrix}\right.$*

Theorem 10.349. *Let $n\in\mathbb{N}\setminus\{1\}$, X a n-dimensional vectorspace over a field F of characterization zero with basis $\{e_i\}_{i\in\{1,\ldots,n\}}$, $Y=S(\{e_i|i\in\{1,\ldots,n-1\}\})$ (see 10.128) a $(n-1)$-dimensional subvector space of X then if $L:Y\rightarrow Y$ is elementair transformation on X then $L_{\langle n\rangle}:X\rightarrow X$ defined by $L_{\langle n\rangle}(e_k)=\left\{\begin{smallmatrix} L(e_k) \text{ if } k\in\{1,\ldots,n-1\}\\ e_n \text{ if } k=n\end{smallmatrix}\right.$ is a elementair trasnformation on Y.*

Proof. As L is a elementair transformation it is of one of the following forms

1. $L=\sigma_{n-1}(i,\alpha)$, $i\in\{1,\ldots,n-1\}$, $\alpha\in F$ then if $k\in\{1,\ldots,n\}$ we have $L_{\langle n\rangle}(e_k)=\left\{\begin{smallmatrix} L(e_k) \text{ if } k\in\{1,\ldots,n-1\}\\ e_n \text{ if } k=n\end{smallmatrix}\right.=\left\{\begin{smallmatrix} \sigma_{n-1}(i,j)(e_k) \text{ if } k\in\{1,\ldots,n-1\}\\ e_k \text{ if } k=n\end{smallmatrix}\right.=\left\{\begin{smallmatrix} e_k \text{ if } k\in\{1,\ldots,n-1\}\setminus\{i\}\\ \alpha\cdot e_i \text{ if } k=i\\ e_k \text{ if } k=n\end{smallmatrix}\right.=\left\{\begin{smallmatrix} e_k \text{ if } k\in\{1,\ldots,n\}\setminus\{i\}\\ \alpha\cdot e_i \text{ if } k=i\end{smallmatrix}\right.=\sigma_n(i,\alpha)(e_k)$, which proves that $L_{\langle n\rangle}=\sigma_n(i,j)$
2. $L=\tau_{n-1}(i,j)$, $i,j\in\{1,\ldots,n-1\}$ then if $k\in\{1,\ldots,n\}$ we have $L_{\langle n\rangle}(e_k)=\left\{\begin{smallmatrix} L(e_k) \text{ if } k\in\{1,\ldots,n-1\}\\ e_n \text{ if } k=n\end{smallmatrix}\right.=\left\{\begin{smallmatrix} \tau_{n-1}(i,j)(e_k) \text{ if } k\in\{1,\ldots,n-1\}\\ e_k \text{ if } k=n\end{smallmatrix}\right.=\left\{\begin{smallmatrix} e_k \text{ if } k\in\{1,\ldots,n-1\}\setminus\{i,j\}\\ e_i \text{ if } k=j\\ e_j \text{ if } k=i\\ e_k \text{ if } k=n\end{smallmatrix}\right.=\left\{\begin{smallmatrix} e_k \text{ if } k\in\{1,\ldots,n\}\setminus\{i.j\}\\ e_i \text{ if } k=j\\ e_j \text{ if } k=i\end{smallmatrix}\right.=\tau_n(i,j)(e_k)$, which proves that $L_{\langle n\rangle}=\tau_n(i,j)$

3. $L=\beta_{n-1}(i,j,\alpha)$, $i,j\in\{1,\ldots,n-1\}$, $\alpha\in F$ then if $k\in\{1,\ldots,n\}$ we have $L_{\langle n\rangle}(e_k)=\begin{cases} L(e_k) \text{ if } k\in\{1,\ldots,n-1\} \\ e_n \text{ if } k=n\end{cases}=\begin{cases}\beta_n(i,j) \text{ if } k\in\{1,\ldots,n-1\} \\ e_k \text{ if } k=n\end{cases}=$ $\begin{cases} e_k \text{ if } k\in\{1,\ldots,n-1\}\setminus\{i\} \\ e_i+\alpha\cdot e_j \text{ if } k=i \\ e_k \text{ if } k=n\end{cases}=\begin{cases} e_k \text{ if } k\in\{1,\ldots,n\}\setminus\{i\} \\ e_i+\alpha\cdot e_j \text{ if } k=i\end{cases}=\beta_n(i,j,\alpha)(e_k)$ which proves that $L_{\langle n\rangle}=\beta_n(i,j,\alpha)$ □

Corollary 10.350. *Let $n\in\mathbb{N}\setminus\{1\}$, X a n-dimensional vector space over a field F of characterization zero with basis $\{e_i\}_{i\in\{1,\ldots,n\}}$, $Y=S(\{e_i|i\in\{1,\ldots,n\}\})$ (see 10.128) a $(n-1)$-dimensional subvector space of X then if $L=L_m\circ\ldots\circ L_1$ where $\{L_i\}_{i\in\{1,\ldots,n\}}$ a familie of elementair transformations on Y then we have $L_{\langle n\rangle}=(L_m)_{\langle n\rangle}\circ\ldots\circ(L_1)_{\langle n\rangle}$ also a composition of elementair transformations (by the above theorem).*

Proof. This is proved by induction on m. so let $S=\{m\in\mathbb{N}|\text{If } L=L_m\circ\ldots\circ L_1$ is the composition of m elementair transformations on Y then $L_{\langle n\rangle}=((L_m)_{\langle n\rangle}\circ\ldots\circ(L_1)_{\langle n\rangle})\}$ then we have

$1\in S$. If $L=L_1\circ\ldots\circ L_1=L_1$ then $L_{\langle n\rangle}=(L_1)_{\langle n\rangle}=(L_1)_{\langle n\rangle}\circ\ldots\circ(L_1)_{\langle n\rangle}$

$m\in S\Rightarrow m+1\in S$. First if $L=L_2\circ L_1$ then $\forall i\in\{1,\ldots,n\}$ we have $L_{\langle n\rangle}(e_i)=\begin{cases} L(e_i) \text{ if } i\in\{1,\ldots,n-1\} \\ e_n \text{ if } i=n\end{cases}=\begin{cases} L_2(L_1(e_i)) \text{ if } i\in\{1,\ldots,n-1\} \\ e_n \text{ if } i=n \text{ if } i=n\end{cases}=$ $\begin{cases}(L_2)_{\langle n\rangle}((L_1)_{\langle n\rangle}(e_i)) \text{ if } i\in\{1,\ldots,n-1\} \\ (L_2)_{\langle n\rangle}((L_1)_{\langle n\rangle}(e_i)) \text{ if } i=n\end{cases}=(L_2)_{\langle n\rangle}((L_1)_{\langle n\rangle}(e_i))$ proving that

$$(L_2\circ L_1)_{\langle n\rangle}=(L_2)_{\langle n\rangle}\circ(L_1)_{\langle n\rangle} \tag{10.113}$$

Let now $L=L_{m+1}\circ\ldots\circ L_1$ then $L_{\langle n\rangle}=(L_{m+1}\circ\ldots\circ L_1)_{\langle n\rangle}=(L_{m+1}\circ(L_m\circ\ldots\circ L_1))_{\langle n\rangle}\underset{10.113}{=}(L_{m+1})_{\langle n\rangle}\circ(L_m\circ\ldots\circ L_1)_{\langle n\rangle}\underset{m\in S}{=}(L_{m+1})_{\langle n\rangle}\circ((L_m)_{\langle n\rangle}\circ\ldots\circ(L_1)_{\langle n\rangle})=((L_{m+1})_{\langle n\rangle}\circ\ldots(L_1)_{\langle n\rangle})$ proving that $m+1\in S$ □

Using the composition of n function on linear transformations we have the following theorem

Lemma 10.351. *Let $n,m\in\mathbb{N}$, X a n-dimensional vector space over a field of characterization zero, $\{L_i\}_{i\in\{1,\ldots.,n\}}$ a family of linear transformations $L_i: X\rightarrow X$ then we have*

1. *$L_m\circ\ldots\circ L_1: X\rightarrow X$ is a linear transformation*
2. $\det(L_m\circ\ldots\circ L_1)=\prod_{i\in\{1,\ldots n\}}\det(L_i)$
3. *If L is regular then $\forall i\in\{1,\ldots,m\}$ we have that L_i is regular (thus a isomorphism because of 10.212).*

Proof. We use induction to prove this theorem

1. Let $S=\{m\in\mathbb{N}|\text{If }\{L_i\}_{i\in\{1,\ldots,m\}}$ is a family of linear transformations on X then $L_m\circ\ldots.\circ L_1$ is *a* linear transformation on $X\}$ then we have

   $1\in S$. If $\{L_i\}_{i\in\{1\ldots,1\}}$ is a family of linear transformations then $L_1\circ\ldots\circ L_1=L_1: X\rightarrow X$ is a linear transformation proving that $1\in S$

$m\in S\Rightarrow m+1\in S$. If $\{L_i\}_{i\in\{1,\ldots,m+1\}}$ is a family of linear transformations on X then as $m\in S$ we have that $L_m\circ\ldots\circ L_1: X\rightarrow X$ is a linear transformation on X. As $L_{m+1}: X\rightarrow X$ is also a linear transformation on X we have by 10.182 that $L_{m+1}\circ\ldots\circ L_1 = L_{m+1}\circ(L_m\circ\ldots\circ L_1): X\rightarrow X$ is a linear transformation on X, proving that $m+1\in S$.

2. Let $S=\left\{m\in\mathbb{N}|\text{If }\{L_i\}_{i\in\{1,\ldots,m\}}\text{ is a family of linear transformations on }X\text{ then }\det(L_m\circ\ldots\circ L_1)=\prod_{i\in\{1,\ldots,m\}}X_i\right\}$ then we have

   $1\in S$. If $\{L_i\}_{i\in\{1\ldots,1\}}$ is a family of linear transformations then $\det(L_1\circ\ldots\circ L_1)=\det(L_1)=\prod_{i\in\{1\ldots,1\}}\det(L_i)$ proving that $1\in S$

   $m\in S\Rightarrow m+1\in S$. If $\{L_i\}_{i\in\{1,\ldots,m+1\}}$ is a family of linear transformations on X then $\det(L_{m+1}\circ\ldots\circ L_1)=\det(L_{m+1}\circ(L_m\circ\ldots\circ L_1))\underset{10.280}{=}\det(L_{m+1})\cdot\det(L_m\circ\ldots\circ L_1)\underset{m\in S}{=}\det(L_{m+1})\cdot\prod_{i\in\{1,\ldots,m\}}\det(L_i)=\prod_{i\in\{1,\ldots,m+1\}}\det(L_i)$

3. Assume that $\{L_i\}_{i\in\{1,\ldots,m\}}$ is a family of linear transformations then by 10.280 we have that $0\neq\det(L_m\circ\ldots\circ L_1)=\prod_{i\in\{1,\ldots,m\}}\det(L_i)$ so that $\forall i\in\{1,\ldots,m\}$ we must have that $\det(L_i)\neq 0$ and thus by 10.280 again we have that L_i is a regular linear transformation. □

Next we need the following two lemmas

Lemma 10.352. *Let $n\in\mathbb{N}$, X a $(n+1)$-dimensional vector space over a field F of characterization zero with basis $\{e_i\}_{i\in\{1,\ldots,n+1\}}$ then if $\{\alpha_i\}_{i\in\{1,\ldots,n\}}$ is a family of values in F then $\forall k\in\{1,\ldots,n\}$ we have $(\beta_{n+1}(n+1,n,\alpha_n)\circ\ldots\circ\beta_{n+1}(n+1,1,\alpha_n))(e_k)=e_k$*

Proof. Let $k\in\{1,\ldots,n\}$ take then $S_k=\{m\in\mathbb{N}|\text{If } m\leqslant n \text{ then } (\beta_{n+1}(n+1,m,\alpha_m)\circ\ldots\circ\beta_{n+1}(n+1,1,\alpha_1))(e_k)=e_k\}$ then we have

$1\in S$. $(\beta_{n+1}(n+1,1,\alpha_1)\circ\ldots\circ\beta_{n+1}(n+1,1,\alpha_1))(e_k)=\beta_{n+1}(n+1,1,\alpha_1)(k)\underset{k\neq n+1}{=}e_k$ proving that $1\in S_k$

$m\in S_k$. Take $m+1$ and assume that $m+1\leqslant n$ then we have $(\beta_{n+1}(n+1,m+1,\alpha_{m+1})\circ\ldots\circ\beta_{n+1}(n+1,1,\alpha_1))(x_k)=(\beta_{n+1}(n+1,m+1,\alpha_{m+1})\circ(\beta_{n+1}(n+1,m,\alpha_m)\circ\ldots\circ\beta_{n+1}(n+1,1,\alpha_1)))(x_k)=\beta_{n+1}(n+1,m+1,\alpha_{m+1})((\beta_{n+1}(n+1,m,\alpha_m)\circ\ldots\circ\beta_{n+1}(n+1,1,\alpha_1))(x_k))\underset{m\in S_k,m\leqslant n}{=}\beta_{n+1}(n+1,m+1,\alpha_{m+1})(x_k)\underset{k\neq n+1}{=}x_k$ proving that $m+1\in S_k$

By mathematical induction we have then that $S_k=\mathbb{N}$, as $n\in\mathbb{N}\Rightarrow n\in S_k$ and $n\leqslant n$, we have $\forall k\in\{1,\ldots,n\}$ that $(\beta_{n+1}(n+1,n,\alpha_n)\circ\ldots\circ\beta_{n+1}(n+1,1,\alpha_n))(e_k)=e_k$ □

Lemma 10.353. *Let $n\in\mathbb{N}$, X a $(n+1)$-dimensional vector space over a field F of characterization zero with basis $\{e_i\}_{i\in\{1,\ldots,n+1\}}$ and $\{\alpha_i\}_{i\in\{1,\ldots,n\}}$ a family of values in F then $(\beta_{N+1}(n+1,n,\alpha_n)\circ\ldots\circ\beta_{n+1}(n+1,1,\alpha_1))(e_{n+1})=(\sum_{k=1}^n\alpha_k\cdot e_k)+e_{n+1}$*

Proof. We prove this by induction so let $S=\{m\in\mathbb{N}|\text{If } m\leqslant n \text{ then } (\beta_{n+1}(n+1, m,\alpha_m)\circ\ldots\circ\beta_{n+1}(n+1,1,\alpha_1))(e_{n+1})=(\sum_{k=1}^m \alpha_k\cdot e_k)+e_{n+1}\}$ then we have

$1\in S$. $(\beta_{n+1}(n+1,1,\alpha_1)\circ\ldots\beta_{n+1}(n+1,1,\alpha_1))(e_{n+1})=\beta_{n+1}(n+1,1, \alpha_1)(e_{n+1})=\alpha_1\cdot e_1+e_{n+1}=(\sum_{i=1}^1\alpha_k\cdot e_k)+e_{n+1}$ proving that $1\in S$

$m\in S\Rightarrow m+1\in S$. Take $m+1$ and assume that $m+1\leqslant n$ then $(\beta_{n+1}(n+1,m+1,\alpha_{m+1})\circ\ldots\circ\beta_{n+1}(n+1,1,\alpha_1))(e_{n+1})=(\beta_{n+1}(n+1,m+1,\alpha_{m+1})\circ(\beta_{n+1}(n+1,m,\alpha_m)\circ\ldots\circ\beta_{n+1}(n+1,1,\alpha_1)))(e_{n+1})=\beta_{n+1}(n+1,m+1,\alpha_{m+1})((\beta_{n+1}(n+1,m,\alpha_m)\circ\ldots\circ\beta_{n+1}(n+1,1,\alpha_1))(e_{n+1}))\underset{m\leqslant n\wedge m\in S}{=}\beta_{n+1}(n+1,m+1,\alpha_{m+1})((\sum_{k=1}^m\alpha_k\cdot e_k)+e_{n+1})=(\sum_{k=1}^m\alpha_k\cdot\beta_{n+1}(n+1,m+1,\alpha_{m+1})(e_k))+\beta_{n+1}(n+1,m+1,\alpha_{m+1})(e_{n+1})=(\sum_{k=1}^m\alpha_k\cdot e_k)+\alpha_{m+1}\cdot e_{m+1}+e_{n+1}=(\sum_{k=1}^{m+1}\alpha_k\cdot e_k)+e_{n+1}$ proving that $m+1\in S$

By mathematical induction we have then that $S=\mathbb{N}$, so as $n\in\mathbb{N}\Rightarrow n\in S$ and $n\leqslant n$ we have $(\beta_{n+1}(n+1,n,\alpha_n)\circ\ldots\circ\beta_{n+1}(n+1,1,\alpha_1))=(\sum_{k=1}^n\alpha_k\cdot e_k)+e_{n+1}$ □

Lemma 10.354. *Let $n\in\mathbb{N}$, X a $(n+1)$-dimensional vector space over a field F of characterization zero with basis $\{e_i\}_{i\in\{1,\ldots,n+1\}}$ and $\{a_i\}_{i\in\{1,\ldots,n\}}$ a family in F then if $m\in\mathbb{N}$ with $m\leqslant n$ is such that $i\in\{m+1,\ldots,n+1\}$ we have that $(\beta_{n+1}(m,n+1,\alpha_m)\circ\ldots\circ\beta_{n+1}(1,n+1,\alpha_1))(e_i)=e_i$*

Proof. This is proved by induction so let $S=\{m\in\mathbb{N}|\text{If } m\leqslant n \text{ then if } i\in\{m+1,\ldots, n+1\} \text{ we have } (\beta_{n+1}(m,n+1,\alpha_m)\circ\ldots\circ\beta_{n+1}(1,n+1))(e_i)=e_i\}$ then we have

$1\in S$. If $i\in\{2,\ldots,n+1\}$ then we have $i\neq 1$ so that $(\beta_{n+1}(1,n+1,\alpha_1)\circ\ldots\circ\beta_{n+1}(1,n+1,\alpha_1))(e_i)=\beta_{n+1}(1,n+1,\alpha_1)(e_i)=e_i$ proving that $1\in S$

$m\in S\Rightarrow m+1\in S$. If $m+1\leqslant n$ then if $i\in\{(m+1)+1,\ldots,n+1\}$ we have $i\in\{(m+1),\ldots,n+1\}$ and $(\beta_{n+1}(m+1,n+1,\alpha_{m+1})\circ\ldots\circ\beta_{n+1}(1,n+1,\alpha_1))(e_1)=\beta_{n+1}(m+1,n+1,\alpha_{m+1})((\beta_{n+1}(m,n+1,\alpha_m)\circ\ldots\circ\beta_{n+1}(1,n+1,\alpha_1))(e_i))\underset{m\in S\wedge i\in\{(m+1),\ldots,n+1\}}{=}\beta_{n+1}(m+1,n+1,\alpha_{m+1})(e_i)\underset{i\neq m+1}{=}e_i$ proving that $m+1\in S$

By mathematical induction we have then that $S=\mathbb{N}$ so if $m\in\mathbb{N}=S$ with $m\leqslant n$ we have if $i\in\{m+1,\ldots,n+1\}$ that $(\beta_{n+1}(m,n+1,\alpha_m)\circ\ldots\circ\beta_{n+1}(1,n+1,\alpha_1))(e_i)=e_i$. □

Lemma 10.355. *Let $n\in\mathbb{N}$, X a $(n+1)$-dimensional vector space over a field F of characterization zero with basis $\{e_i\}_{i\in\{1,\ldots,n+1\}}$ and $\{a_i\}_{i\in\{1,\ldots,n\}}$ a matrix in F then given $i\in\{1,\ldots,n\}$ we havc $(\beta_{n+1}(n,n+1,\alpha_{n,})\circ\ldots\circ\beta_{n+1}(1,n+1,\alpha_i))(e_i)=e_i+\alpha_i\cdot e_{n+1}$*

Proof. We prove this by induction so let $S_i=\{m\in\mathbb{N}|\text{If } m\leqslant n \text{ then if } i\in\{1,\ldots,m\}$ we have $(\beta_{n+1}(m,n+1,\alpha_m)\circ\ldots\circ\beta_{n+1}(1,n+1,\alpha_1))(e_i)=e_i+\alpha_i\cdot e_{n+1}\}$ then we have

$1\in S_i$. For if $m=1$ then if $i\in\{1,\ldots,1\}\Rightarrow i=1$ we have $(\beta_{n+1}(1,n+1,\alpha_1)\circ\ldots\circ\beta_{n+1}(1,n+1,\alpha_1))(e_i)=\beta_{n+1}(1,n+1,a_1)(e_1)=e_1+\alpha_i\cdot e_{n+1}$ proving that $1\in S_i$

$m\in S\Rightarrow m+1\in S$. Assume that $m+1\leqslant n$ and that $i\in\{1,\ldots,m+1\}$ then we have either

$i\in\{1,\ldots,m\}$. then $(\beta_{n+1}(m+1,n+1,\alpha_{m+1})\circ\ldots\circ\beta_{n+1}(1,n+1,\alpha_i))(e_i)=\beta_{n+1}(m+1,n+1,\alpha_{m+1})((\beta_{n+1}(m,n+1,\alpha_m)\circ\ldots\circ\beta_{n+1}(1,n+1,\alpha_1))(e_i)) \underset{m\in S\wedge i\in\{1,\ldots,m\}\wedge m\leqslant n}{=} \beta_{n+1}(m+1,n+1,\alpha_{m+1})(e_i+\alpha_i\cdot e_{n+1})=\beta_{n+1}(m+1,n+1,\alpha_{m+1})(e_i)+\alpha_i\cdot\beta_{n+1}(m+1,n+1,\alpha_{m+1})(e_{n+1}) \underset{i,n+1\neq m+1}{=} e_i+\alpha_i\cdot e_{n+1}$

$i=m+1$. then $(\beta_{n+1}(m+1,n+1,\alpha_{m+1})\circ\ldots\circ\beta_{n+1}(1,n+1,\alpha_i))(e_i)=\beta_{n+1}(m+1,n+1,\alpha_{m+1})((\beta_{n+1}(m,n+1,\alpha_m)\circ\ldots\circ\beta_{n+1}(1,n+1,\alpha_1))(e_i)) \underset{\text{Lemma }10.354}{=} \beta_{n+1}(m+1,n+1,\alpha_{m+1})(e_i)=e_{m+1}+\alpha_{m+1}\cdot e_{n+1}=e_i+\alpha_i\cdot e_{n+1}$

This proves that $m+1\in S$

Using mathematical induction we have $\mathbb{N}=S$ so as $n\in\mathbb{N}=S$ and $n\leqslant n$ we have $(\beta_{n+1}(n,n+1,\alpha_{n,})\circ\ldots\circ\beta_{n+1}(1,n+1,\alpha_i))(e_i)=e_i+\alpha_i\cdot e_{n+1}$ □

Theorem 10.356. *Let $n\in\mathbb{N}$, X a n-dimensional space over a field F of characterization zero, $L\in\mathrm{Hom}(X,X)$ a regular linear transformation (hence a isomorphism) then there exists a family $\{L_i\}_{i\in\{1,\ldots,m\}}$ of elementair regular transformations so that $L=L_n\circ\ldots\circ L_1$*

Proof. We prove this by induction on n so let $S=\{n\in\mathbb{N}|\text{If } X \text{ is a n-dimensional vector space over } F\text{ , } L\in\mathrm{Hom}(X,X) \text{ a regular transformation then } \exists\{L_i\}_{i\in\{1,\ldots,m\}} \text{ of regular elementair transformations so that } L=L_m\circ\ldots\circ L_1\}$, we have then

$1\in S$. If X is one dimensional it has a base $\{e_i\}_{i\in\{1,\ldots,1\}}$ then if $L\in\mathrm{Hom}(X,X)$ we have $L(e_1)=\alpha\cdot e_1$ so $L=\sigma_1(1,\alpha)$ a elementair function, which must be regular as L is regular, from $L=\sigma_1(1,\alpha)=\sigma_1(1,\alpha)\circ\ldots\circ\sigma_1(1,\alpha)$ we have then that $1\in S$.

$n\in S\Rightarrow n+1\in S$. Take now X a $(n+1)$-dimensional space and $L\in\mathrm{Hom}(X,X)$ a regular transformation. Define now $M'=\{M'_{i,j}\}_{i,j\in\{1,\ldots,n+1\}\times\{1,\ldots,n+1\}}=\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n+1\}},\{e_i\}_{i\in\{1,\ldots,n+1\}})$ the associated matrix of L so that L is defined by

$$\forall i\in\{1,\ldots,n\}\vDash L(e_i)=\sum_{k=1}^{n+1} M'_{k,i}e_k \tag{10.114}$$

As L is regular we have that $0\neq\det(L)\underset{10.319}{=}\det(M')\underset{10.340,j=n+1}{=}\sum_{i\in\{1,\ldots,n+1\}}(-1)^{i+(n+1)}M'_{i,j}\cdot\det([i\boxplus(n+1)]M')$ so that $\exists i_0\in\{1,\ldots,n+1\}$ so that

$$\det(i_0\boxplus(n+1)M')\neq 0 \tag{10.115}$$

We have now for i_0 the following possibilities

$i_0 = n+1$. So $\det([n+1 \boxplus n+1]M) \neq 0$, take then $T = 1_{\{1,\ldots,n+1\}} = \tau_{n+1}(n+1, n+1)$ a elementair transformation and $L_1 = L$ then we have $L = T \circ L$ and

There exists a elementair transformation T such that $L = T \circ L_1$, $M = \mathcal{M}(L_1, \{e_i\}_{i \in \{1,\ldots,n+1\}}, \{e_i\}_{i \in \{1,\ldots,n+1\}}), \det([n+1 \boxplus n+1]M) \neq 0$ (10.116)

$i_0 \leqslant n$. Consider now $([i_0 \boxplus (n+1)]M')_{k,l}$ then we have the following cases (based on the definition) for $k, l \in \{1, \ldots, n\}$

$k < i_0 \wedge l < n+1$. then $([i_0 \boxplus (n+1)]M')_{k,l} = M'_{k,l}$

$i_0 \leqslant k \leqslant n \wedge l < n+1$. then $([i_0 \boxplus (n+1)]M')_{k,l} = M'_{k+1,l}$

$k < i_0 \wedge n+1 \leqslant l \leqslant n$. does not occur as $n < n+1$

$i_0 \leqslant k \leqslant n \wedge n+1 \leqslant l \leqslant n$. does not occur as $n < n+1$

proving that

$\forall k, l \in \{1, \ldots, n\}$ we have $([i_0 \boxplus (n+1)]M')_{k,l} = \begin{cases} M'_{k,l} \text{ if } 1 \leqslant k < i_0 \\ M'_{k+1,l} \text{ if } i_0 \leqslant k \leqslant n \end{cases}$ (10.117)

Take now $(n \rightsquigarrow_n i_0): \{1, \ldots, n\} \rightarrow \{1, \ldots, n\}$ (see 10.93 for the definition) then we have for $([i_0 \boxplus (n+1)]M')_{(n \rightsquigarrow_n i_0)}$ the following cases

$i_0 = n$. then $\forall k, l \in \{1, \ldots, n\}$ we have $(([i_0 \boxplus (n+1)]M')_{(n \rightsquigarrow_n n)})_{k,l} = ([i \boxplus (n+1)]M')_{k,l} = \begin{cases} M'_{k,l} \text{ if } 1 \leqslant k < n \\ M'_{n+1,l} \text{ if } k = n \end{cases} = \begin{cases} M'_{k,l} \text{ if } k \in \{1, \ldots, n\} \setminus \{i_o\} \\ M'_{n+1,l} \text{ if } k = i_0 \end{cases}$

$i_0 < n$. then $\forall k, l \in \{1, \ldots, n\}$ we have

$$\begin{aligned}
(([i_0 \boxplus (n+1)]M')_{(n \rightsquigarrow_n i_0)})_{k,l} &= ([i_0 \boxplus (n+1)]M')_{(n \rightsquigarrow_n i_0)(k),l} \\
&= \begin{cases} ([i_0 \boxplus (n+1)]M')_{k,l} \text{ if } k < i_0 \\ ([i_0 \boxplus (n+1)]M')_{n,l} \text{ if } k = i_0 \\ ([i_0 \boxplus (n+1)]M')_{k-1,l} \text{ if } i_0 < k \leqslant n \\ ([i_0 \boxplus (n+1)]M')_{k,l} \text{ if } n < k \text{ (impos.)} \end{cases} \\
&\underset{10.117}{=} \begin{cases} M'_{k,l} \text{ if } k < i_0 \\ M'_{n+1} \text{ if } k = i_0 \\ M'_{(k-1)+1,l} \text{ if } i_0 < k \leqslant n \end{cases}
\end{aligned}$$

$$= \begin{cases} M'_{k,l} \text{ if } k < i_0 \\ M'_{n+1,l} \text{ if } k = i_0 \\ M'_{k,l} \text{ if } i_0 < k \leqslant n \end{cases}$$

$$= \begin{cases} M'_{k,l} \text{ if } k \in \{1, \ldots, n\} \setminus \{i_0\} \\ M'_{n+1,l} \text{ if } k = i_0 \end{cases}$$

So we have that in all cases that $\forall k, l \in \{1, \ldots, n\}$ we have $(([i_0 \boxplus (n+1)]M_i)_{(n \rightsquigarrow_n i_0)})_{k,l} = \begin{cases} M'_{k,l} \text{ if } k \neq i_0 \\ M'_{n+1,l} \text{ if } k = i_0 \end{cases}$ so we have

If M'' is defined by $\forall k, l \in \{1, \ldots, n\} \vDash M''_{k,l} = \begin{cases} M'_{k,l} \text{ if } k \neq i_0 \\ M'_{n+1,l \text{ if } k = i_0} \end{cases}$ then $M'' = ([i_0 \boxplus (n+1)]M')_{(n \rightsquigarrow i_0)}$ (10.118)

Now we have that $\det(M'') = \det(([i_0 \boxplus (n+1)M'])_{n \rightsquigarrow_n i_0}) \underset{10.323}{=} \operatorname{sign}(n \rightsquigarrow_n i_0) \cdot \det([i_0 \boxplus (n+1)]M') \neq 0$ (using 10.115) proving that

$$\det(M'') \neq 0 \tag{10.119}$$

Define now $L_1 = \tau_{n+1}(n+1, i_0) \circ L$ and let $M = \mathcal{M}(L_1, \{e_i\}_{i \in \{1, \ldots, n+1\}}, \{e_i\}_{i \in \{1, \ldots, n+1\}})$ then we have that $\forall i \in \{1, \ldots, n\}$

$$\begin{aligned} L_1(e_i) &= \tau_{n+1}(n+1, i_0)(L(e_i)) \\ &= \tau_{n+1}(n+1, i_0)\left(\sum_{k \in \{1, \ldots, n\}} M'_{k,i} \cdot e_k\right) \\ &= \sum_{k \in \{1, \ldots, n\}} M'_{k,i} \cdot \tau_{n+1}(n+1, i_0)(e_k) \\ &= \sum_{k \in \{1, \ldots, n\} \setminus \{n+1, i_0\}} M'_{k,i} \cdot \tau_{n+1}(n+1, i_0)(e_k) + M'_{i_0,i} \cdot \\ &\quad \tau_{n+1}(n+1, i_0)(e_{i_0}) + M'_{n+1,i} \cdot \tau_{n+1}(n+1, i_0)(e_{n+1}) \\ &= \sum_{k \in \{1, \ldots, n\} \setminus \{n+1, i_0\}} M'_{k,i} \cdot e_k + M'_{i_0,i} \cdot e_{n+1} + M'_{n+1,i} \cdot e_{i_0} \\ &= \sum_{k \in \{1, \ldots, n\} \setminus \{n+1, i_0\}} M_{k,i} \cdot e_k + M_{i_0,i} \cdot e_{n+1} + M_{n+1,i} \cdot e_{i_0} \end{aligned}$$

proving by the uniqueness of the expansion in a basis that M is defined by $\forall k, l \in \{1, \ldots, n+1\}$ we have $M_{k,l} = \begin{cases} M_{k,l} \text{ if } k \in \{1, \ldots n\} \setminus \{n+1, i_0\} \\ M_{n+1,l} \text{ if } k = i_0 \\ M_{i_0,l} \text{ if } k = n+1 \end{cases}$.

Now $\forall k, l \in \{1, \ldots, n\}$ we have

$$\begin{aligned} ([n+1 \boxplus n+1]M)_{k,l} &\underset{k,l<n+1}{=} M_{k,l} \\ &\underset{i_0 \leqslant n < n+1}{=} \begin{cases} M_{k,l} \text{ if } k \in \{1, \ldots, n\} \setminus \{i_0\} \\ M_{n+1} \text{ if } k = i_0 \end{cases} \\ &= M'' \end{aligned}$$

Using 10.119 we have then that $\det([n+1 \boxplus n+1]M) = \det(M'') \neq 0$. As $L_1 = \tau_{n+1}(n+1, i_0) \circ L \underset{10.346}{\Rightarrow} L = \tau_{n+1}(n+1, i_0) \circ L_1$. Take $T = \tau_{n+1}(n+1, i_0)$ a elementair transformation then we have

There exist a elementair transformation T such that $L = T \circ L_1$, $M = \mathcal{M}(L_1, \{e_i\}_{i\in\{1,\ldots,n+1\}}, \{e_i\}_{i\in\{1,\ldots,n+1\}})$, $\det([n+1, n+1]M)$ (10.120)

Using 10.116 and 10.120 we have then that

$$\exists T \text{ a elementair transformation such that } L = T \circ L_1 \tag{10.121}$$

and

If $M = \mathcal{M}(L_1, \{e_1\}_{i\in\{1,\ldots,n+1\}}, \{e_i\}_{i\in\{1,\ldots,n+1\}})$ then $\det([n+1 \boxplus n+1]M) \neq 0$ (10.122)

Take now $Y = \mathcal{S}(\{e_1, \ldots, e_n\})$ (see 10.128) a n-dimensional subspace of X take then

$$L_2' \in \operatorname{Hom}(Y, Y) \text{ defined by } \forall i \in \{1, \ldots, n\} \vDash L_2'(e_i) = \sum_{k\in\{1,\ldots,n\}} M_{k,i} \cdot e_k \tag{10.123}$$

then we have $\det(L_2') = \det([n+1 \boxplus n+1]M) \neq 0$ so as $n \in S$ there exists a family of elementair transformations $\{T_i'\}_{i\in\{1,\ldots,m\}}$ such that $L_2' = T_m' \circ \ldots \circ T_1'$. If we take then $L_2 = (L_2)_{\langle n+1\rangle}$ and $\forall i \in \{1, \ldots, m\}$ $T_i = (T_i')_{\langle n+1\rangle}$ then using 10.350 we have $L_2 = T_m \circ \ldots \circ T_1$ giving

If L_2 defined by $\forall i \in \{1, \ldots, n+1\} \vDash L_2(e_i) = \begin{cases} \sum_{k\in\{1,\ldots,n\}} M_{k,i} \cdot e_k \\ L(e_{n+1}) = e_{n+1} \end{cases}$ then there exists a family $\{T_i\}_{i\in\{1,\ldots,m\}}$ of elementair functions in X such that $L_2 = T_m \circ \ldots \circ T_1$ (10.124)

Define now L_3 by

$$L_3 = (\beta_{n+1}(n+1, n, M_{n,n+1}) \circ \ldots \circ \beta_{n+1}(n+1, 1, M_{1,n+1})) \circ L_2 \tag{10.125}$$

then we have for $i \in \{1, \ldots, n\}$ that

$$\begin{aligned} L_3(e_i) &= (\beta_{n+1}(n+1, n, M_{n,n+1}) \circ \ldots \circ \beta_{n+1}(n+1, 1, M_{1,n+1}))(L_2(e_i)) \\ &= (\beta_{n+1}(n+1, n, M_{n,n+1}) \circ \ldots \circ \beta_{n+1}(n+1, 1, M_{1,n+1}))\left(\sum_{k\in\{1,\ldots,n\}} M_{k.i} \cdot e_i\right) \\ &= \sum_{k\in\{1,\ldots,n\}} M_{k,i} \cdot (\beta_{n+1}(n+1, n, M_{n,n+1}) \circ \ldots \circ \beta_{n+1}(n+1, 1, M_{1,n+1}))(e_i) \\ &\underset{10.352}{=} \sum_{k\in\{1,\ldots,n\}} M_{k,i} \cdot e_i \end{aligned}$$

and for $n+1$ we have

$$\begin{array}{rcl}L_3(e_{n+1}) & = & (\beta_{n+1}(n+1,n,M_{n,n+1})\circ\ldots\circ\beta_{n+1}(n+1,1,M_{1,n+1}))(e_{n+1})\\ & \underset{10.353}{=} & \sum_{k\in\{1,\ldots,n\}} M_{k,n+1}\cdot e_k+e_{n+1}\end{array}$$

which proves that

$\forall i \in \{1, \ldots, n+1\}$ we have $L_3(e_i) =$

$$\begin{cases}\sum_{k\in\{1,\ldots,n\}} M_{k,i}\cdot e_k \text{ if } i\in\{1,\ldots,n\}\\ \left(\sum_{k\in\{1,\ldots,n\}} M_{k,n+1}\cdot e_k\right)+e_{n+1}\end{cases} \tag{10.126}$$

Next take $\alpha\in\mathbb{R}$ define then

$$L_4=\sigma_{n+1}(n+1,\alpha)\circ L_3 \tag{10.127}$$

Then we have $\forall i\in\{1,\ldots,n\}$ that

$$\begin{array}{rcl}L_4(e_i) & = & \sigma_{n+1}(n+1,\alpha)(L_3(e_i))\\ & \underset{10.126}{=} & \sigma_{n+1}(n+1,\alpha)\left(\sum_{k\in\{1,\ldots,n\}} M_{k,i}\cdot e_k\right)\\ & = & \sum_{k\in\{1,\ldots,n\}} M_{k,i}\cdot\sigma_{n+1}(n+1,\alpha)(e_k)\\ & = & \sum_{k\in\{1,\ldots,n\}} M_{k,i}\cdot e_k\end{array}$$

and for $n+1$ we have

$$\begin{array}{rcl}L_4(e_{n+1}) & = & \sigma_{n+1}(n+1,\alpha)(L_4(e_{n+1}))\\ & \underset{10.126}{=} & \sigma_{n+1}(n+1,\alpha)\left(\left(\sum_{k\in\{1,\ldots,n\}} M_{k,i}\cdot e_k\right)+e_{n+1}\right)\\ & = & \left(\sum_{k\in\{1,\ldots,n\}} M_{k,i}\cdot\sigma_{n+1}(n+1,\alpha)(e_k)\right)+\sigma_{n+1}(n+1,\\ & & \alpha)(e_{n+1})\\ & = & \sum_{k\in\{1,\ldots,n\}} M_{k,i}\cdot e_k+\alpha\cdot e_{n+1}\end{array}$$

which proves that

$\forall i \in \{1, \ldots, n+1\} \vDash L_4(e_i) =$

$$\begin{cases}\sum_{k\in\{1,\ldots,n\}} M_{k,i}\cdot e_k \text{ if } i\in\{1,\ldots,n\}\\ \left(\sum_{k\in\{1,\ldots,n\}} M_{k,n+1}\cdot e_k\right)+\alpha\cdot e_{n+1} \text{ if } i=n+1\end{cases} \tag{10.128}$$

Next given $B=M(n\times 1,F)$ (we will define B later) and take

$$L_5=(\beta_{n+1}(n,n+1,B_n)\circ\ldots\circ\beta_{n+1}(1,n+1,B_1))\circ L_4 \tag{10.129}$$

then we have $\forall i \in \{1, \ldots, n\}$ that

$$
\begin{aligned}
L_5(e_i) &= (\beta_{n+1}(n, n+1, B_n) \circ \ldots \circ \beta_{n+1}(1, n+1, B_1))(L_4(e_i)) \\
&\underset{10.128}{=} (\beta_{n+1}(n, n+1, B_n) \circ \ldots \circ \beta_{n+1}(1, n+1, B_1))\left(\sum_{k \in \{1, \ldots, n\}} M_{k,i} \cdot e_k\right) \\
&= \sum_{k \in \{1, \ldots, n\}} M_{k,i} \cdot (\beta_{n+1}(n, n+1, B_n) \circ \ldots \circ \beta_{n+1}(1, n+1, B_1))(e_k) \\
&\underset{10.355}{=} \sum_{k \in \{1, \ldots, n\}} M_{k,i} \cdot (e_k + B_k \cdot e_{n+1}) \\
&= \left(\sum_{k \in \{1, \ldots, n\}} M_{k,i} \cdot e_k\right) + \left(\sum_{k \in \{1, \ldots, n\}} M_{k,i} \cdot B_k\right) \cdot e_{n+1}
\end{aligned}
$$

and for $n+1$

$$
\begin{aligned}
L_5(e_{n+1}) &= (\beta_{n+1}(n, n+1, B_n) \circ \ldots \circ \beta_{n+1}(1, n+1, B_1))(L_4(e_{n+1})) \\
&\underset{10.128}{=} (\beta_{n+1}(n, n+1, B_n) \circ \ldots \circ \beta_{n+1}(1, n+1, B_1))\left(\left(\sum_{k \in \{1, \ldots, n\}} M_{k,n+1} \cdot e_k\right) + \alpha \cdot e_{n+1}\right) \\
&= \left(\sum_{k \in \{1, \ldots, n\}} M_{k,n+1} \cdot (\beta_{n+1}(n, n+1, B_n) \circ \ldots \circ \beta_{n+1}(1, n+1, B_1))(\beta_{n+1}(n, n+1, B_n) \circ \ldots \circ \beta_{n+1}(1, n+1, B_1))(e_k)\right) + \\
&\quad \alpha \cdot (\beta_{n+1}(n, n+1, B_n) \circ \ldots \circ \beta_{n+1}(1, n+1, B_1))(e_{n+1}) \\
&\underset{10.355}{=} \left(\sum_{k \in \{1, \ldots, n\}} M_{k,n+1} \cdot (e_k + B_k \cdot e_{n+1})\right) + \alpha \cdot (\beta_{n+1}(n, n+1, B_n) \circ \ldots \circ \beta_{n+1}(1, n+1, B_1))(e_{n+1}) \\
&\underset{10.354}{=} \left(\sum_{k \in \{1, \ldots, n\}} M_{k,n+1} \cdot (e_k + B_k \cdot e_{n+1})\right) + \alpha \cdot e_{n+1} \\
&= \left(\sum_{k \in \{1, \ldots, n\}} M_{k,n+1} \cdot e_k\right) + \left(\alpha + \sum_{k \in \{1, \ldots, n\}} M_{k,n+1} \cdot B_k\right) \cdot e_{n+1}
\end{aligned}
$$

proving that $\forall i \in \{1, \ldots, n+1\}$ we have

$$
L_5(e_i) = \begin{cases} \left(\sum_{k \in \{1, \ldots, n\}} M_{k,i} \cdot e_k\right) + \left(\sum_{k \in \{1, \ldots, n\}} M_{k,i} \cdot B_k\right) \cdot e_{n+1} & \text{if } i \in \{1, \ldots, n\} \\ \left(\sum_{k \in \{1, \ldots, n\}} M_{k,n+1} \cdot e_k\right) + \left(\alpha + \sum_{k \in \{1, \ldots, n\}} M_{k,n+1} \cdot B_k\right) \cdot e_{n+1} & \text{if } i = n+1 \end{cases} \tag{10.1}
$$

Now $\sum_{k \in \{1,\ldots,n\}} M_{k,i} \cdot B_k = \sum_{k \in \{1,\ldots,n\}} M_{i,k}^T \cdot B_k$ is the matrix product $([n+1 \boxplus n+1]M)^T \cdot B$, we want this product to be $A \in M(n \times 1, F)$ where $\forall i \in \{1, \ldots, n\}$ $A_i = M_{n+1,i}$ so that we have the correct factor for e_{n+1} in 10.130, so we must have

$$([n+1 \boxplus n+1]M)^T \cdot B = A \tag{10.131}$$

Now as $\det(([n+1 \boxplus n+1]M)^T) \underset{10.318}{=} \det([n+1 \boxplus n+1]M) \neq 0$ [by using 10.122] there exists a inverse of $([n+1 \boxplus n+1]M)^T$, take then $N = (([n+1 \boxplus n+1]M)^T)^{-1}$ so that

$$\forall i, j \in \{1, \ldots, n\} \text{ we have } \delta_{i,j} = \sum_{k \in \{1,\ldots,n\}} (M^T)_{i,k} \cdot N_{k,j} = \sum_{k \in \{1,\ldots,n\}} M_{k,i} \cdot N_{k,j} \tag{10.132}$$

Multiplying then by N in 10.131 gives $N \cdot A = N \cdot ([n+1 \boxplus n+1]M)^T \cdot B = (([n+1 \boxplus n+1]M)^T)^{-1} \cdot ([n+1 \boxplus n+1]M)^T \cdot B = B \Rightarrow N \cdot A = B$, so we take B to be defined by

$$\forall i \in \{1, \ldots, n\} \vDash B_i = \sum_{l \in \{1,\ldots,n\}} N_{i,l} \cdot M_{n+1,l} \tag{10.133}$$

Then we have $\forall i \in \{1, \ldots, n\}$ that

$$\begin{aligned}
\sum_{k \in \{1,\ldots,n\}} M_{k,i} \cdot B_k &= \sum_{k \in \{1,\ldots,n\}} M_{k,i} \cdot \left(\sum_{l \in \{1,\ldots,n\}} N_{k,l} \cdot M_{n+1,l} \right) \\
&= \sum_{k \in \{1,\ldots,n\}} \left(\sum_{l \in \{1,\ldots,n\}} (M_{k,i} \cdot N_{k,l}) \cdot M_{n+1,l} \right) \\
&\underset{10.48}{=} \sum_{l \in \{1,\ldots,n\}} \left(\sum_{k \in \{1,\ldots,n\}} M_{k,i} \cdot N_{k,l} \right) \cdot M_{n+1,l} \\
&\underset{10.132}{=} \sum_{l \in \{1,\ldots,n\}} \delta_{i,l} \cdot M_{n+1,l} \\
&= M_{n+1,i}
\end{aligned}$$

Using the above in 10.130 gives that

$$\forall i \in \{1, \ldots, n\} \vDash L_5(e_i) = \sum_{k = \{1,\ldots,n+1\}} M_{k,i} \cdot e_k \tag{10.134}$$

Take now $\alpha = M_{n+1,n+1} - \sum_{k \in \{1,\ldots,n\}} M_{k,n+1} \cdot B_k$ then we have using 10.130 that

$$L_5(e_{n+1}) = \sum_{k \in \{1,\ldots,n+1\}} M_{k,n+1} \cdot e_k \tag{10.135}$$

So we have that $\mathcal{M}(L_5, \{e_i\}_{i\in\{1,\ldots,n+1\}}, \{e_i\}_{i\in\{1,\ldots,n+1\}}) = \mathcal{M}(L_1, \{e_i\}_{i\in\{1,\ldots,n+1\}}, \{e_i\}_{i\in\{1,\ldots,n+1\}}) \underset{10.307}{\Rightarrow} L_5 = L_1$ which proves by 10.121, 10.124, 10.125, 10.125, 10.127 and 10.129 that $L = T\circ(\beta_{n+1}(n, n+1, B_n)\circ\ldots\circ\beta_{n+1}(1, n+1, B_1))\circ\sigma_{n+1}(n+1,\alpha)\circ(\beta_{n+1}(n+1, n, M_{n,n+1})\circ\ldots\circ\beta_{n+1}(n+1, 1, M_{1,n+1}))\circ(T_m\circ\ldots\circ T_1)$. Define now $\forall i\in\{1,\ldots,2+m+2\cdot n\}$ the elementair transformation E_i by

$$E_i = \begin{cases} T_i \text{ if } i\in\{1,\ldots,m\} \\ \beta_{n+1}(n+1, i-m, M_{i-m,n+1}) \text{ if } i\in\{m+1,\ldots,m+n\} \\ \sigma_{n+1}(n+1,\alpha) \text{ if } i = m+n+1 \\ \beta_{n+1}(i-(m+n+1), n+1, B_{(i-(m+n+1))}) \text{ if } i\in\{m+n+2,\ldots,m+2\cdot n+1\} \\ T \text{ if } i = m+2\cdot n+2 \end{cases}$$

then we have $L = E_{m+2\cdot n+2}\circ(E_{m+2\cdot n+1}, \ldots, \circ E_{m+n+2})\circ E_{m+n+1}\circ(E_{m+n}\circ\ldots\circ E_{m+1})\circ(E_m\circ\ldots\circ E_1)$, repeately applying the general associativity of the composition of functions (see 4.81) gives then finally

$$L = E_{m+2\cdot n+2}\circ\ldots\circ E_1$$

Using 10.351 we have as L is regular that then also every E_i is a regular transformation, so that we have proved that L is the composition of regular elementair transformations and thus that $n+1\in S$.

Using mathematical induction we have then that $S=\mathbb{N}$ which proves the theorem. □

10.11 Direct internal sum of sub spaces

We extend now the concept of a direct sum introduced in 10.192 to a direct sum of a arbitrary finite number of sub spaces. First we consider sum's of vector spaces.

Definition 10.357. *If V is a vector space over a field F and $\{V_i\}_{i\in I}$ is a **finite** family of non empty sub spaces then $\sum_{i\in I} V_i = \{v\in V\mid\exists\{v_i\}_{i\in I}$ with $\forall i\in I$ we have $v_i\in V_i$ such that $v=\sum_{i\in I} v_i\}$.*

Example 10.358. If V is a vector space over a field F and $\emptyset=\{V_i\}_{i\in\emptyset}$ is the empty family then $\sum_{i\in\emptyset} V_i = \{0\}$

Proof. If $x\in\sum_{i\in\emptyset} V_i$ then there exists a $\{v_i\}_{i\in\emptyset}=\emptyset$, a family where $\forall i\in\emptyset\vDash v_i\in V_i$ is satisfied vacuously where $x=\sum_{i\in\emptyset} v_i$. As $\text{support}(\{v_i\}_{i\in\emptyset})=\emptyset$ we have $\sum_{i\in\emptyset} v_i = 0$ giving that $x=\{0\}$ □

Example 10.359. If V is a vector space then if $\{V_i\}_{i\in\{k\}}$ is a finite family of sub spaces then $\sum_{i\in\{k\}} V_k = V_k$

Proof. If $v\in\sum_{i\in\{k\}}V_k$ then there exists a $\{v_i\}_{i\in\{k\}}$ such that $v=\sum_{i\in\{k\}}v_i=v_k\in V_k$. If $v\in V_k$ then if we take $\{v_i\}_{i\in\{k\}}$ where $v_k=v$ then we have $v=v_k=\sum_{i\in\{k\}}v_i\in\sum_{i\in\{k\}}V_i$ □

As the sum of sub spaces is based on the generalized sum a lot of its properties transfer to the sum of sub spaces. This is expressed in the next theorems.

Theorem 10.360. *If V is vector space over a field F and $\{V_i\}_{i\in I}$ is a **finite** family of non empty sub spaces then if $h: J\rightarrow I$ is a bijection then we have that $\sum_{j\in J}V_{h(j)}=\sum_{i\in I}V_i$*

Proof. If $v\in\sum_{j\in J}V_{h(j)}$ then $v=\sum_{j\in J}v_j$ where $\forall j\in J$ we have $v_j\in V_{h(j)}$ and if we define $\{w_i\}_{i\in I}$ by $w_i=v_{h^{-1}(i)}$ then $w_{h(j)}=v_{h^{-1}(h(j))}=v_j\in V_{h(j)}$ and $v=\sum_{j\in J}w_{h(j)}\underset{10.44}{=}\sum_{i\in I}w_i\in\sum_{i\in I}V_i$ proving that $\sum_{j\in J}V_{h(j)}\subseteq\sum_{i\in I}V_i$.

If $v\in\sum_{i\in I}V_i$ then $v=\sum_{i\in I}v_i$ where $\forall i\in I$ we have $v_i\in V_i$. Define then $\{w_j\}_{j\in J}$ by $\forall j\in J$ $w_j=v_{h(j)}\in V_{h(j)}$ then $w_{h^{-1}(i)}=v_{h(h^{-1}(i))}=v_i$ which as h^{-1} is a bijection we can use 10.44 again to prove that $v=\sum_{i\in I}w_{h^{-1}(i)}=\sum_{j\in J}w_j\in V_{h(j)}$ proving that $\sum_{i\in I}V_i\subseteq\sum_{j\in J}V_{h(j)}$. □

Theorem 10.361. *If V is vector space over a field F and $\{V_i\}_{i\in I}$ is a **finite** family of non empty sub spaces then if $i\in I$ we have that $\sum_{i\in I}V_i=\left(\sum_{j\in I\setminus\{i\}}V_j\right)+V_i$*

Proof. If $v\in\sum_{j\in I}V_j$ then $v=\sum_{j\in I}v_j$ where $\forall j\in I$ we have $v_j\in V_j$. From this it follows that $v=\sum_{j\in I}v_j=\sum_{j\in I\setminus\{i\}}v_j+\sum_{j\in\{i\}}v_j=\sum_{j\in I\setminus\{i\}}v_j+v_i\in\left(\sum_{j\in I\setminus\{i\}}V_j\right)+V_i$ proving that $\sum_{i\in I}V_i\subseteq\left(\sum_{j\in I\setminus\{i\}}V_j\right)+V_i$.

If $v\in\left(\sum_{j\in I\setminus\{i\}}V_j\right)+V_i$ then there exists a $\{v_j\}_{j\in I\setminus\{i\}}$ and a $w\in V_i$ such that $v=\sum_{j\in I\setminus\{i\}}v_i+w$ define now $\{w_j\}_{j\in I}$ by $w_j=\begin{cases}v_j \text{ if } j\in I\setminus\{i\}\\ w \text{ if } j=i\end{cases}$ then $v=\sum_{j\in I\setminus\{i\}}v_i+w=\sum_{j\in I\setminus\{i\}}w_j+w_i=\sum_{j\in I\setminus\{i\}}w_j+\sum_{j\in\{i\}}w_j=\sum_{j\in I}w_j\in\sum_{j\in I}V_j$ proving that $\left(\sum_{j\in I\setminus\{i\}}V_j\right)+V_i\subseteq\sum_{j\in I}V_j$. □

We can even generalize the above theorem

Theorem 10.362. *If V is a vector space over a field F and $\{V_i\}_{i\in I}$ is a **finite** family of non empty sub spaces then if $I=I_1\bigcup I_2$ where $I_1\bigcap I_2$ then we have $\sum_{i\in I}V_i=(\sum_{i\in I_1}V_i)+(\sum_{i\in I_2}V_i)$*

Proof. If $v\in\sum_{j\in I}V_j$ then $v=\sum_{j\in I}v_j$ where $\forall j\in I$ we have $v_j\in V_j$. From this it follows that $v=\sum_{j\in I}v_j=\sum_{j\in I_1}v_j+\sum_{j\in I_2}v_j\in\sum_{j\in I_1}V_j+\sum_{j\in I_2}V_j$ proving that $\sum_{j\in I}V_i\subseteq\sum_{j\in I_1}V_j+\sum_{j\in I_2}V_j$.

If $v\in\sum_{j\in I_1}V_j+\sum_{j\in I_2}V_j$ then $v=u+w$ where $u=\sum_{j\in I_1}u_j$ and $w=\sum_{j\in I_2}w_j$. Define now (using $I_1\bigcap I_2$) $\{v_i\}_{i\in I}$ by $v_i=\begin{cases}u_i \text{ if } i\in I_1\\ w_i \text{ if } i\in I_2\end{cases}$ then $\sum_{j\in I}V_j\ni\sum_{j\in I}v_j=\sum_{j\in I_1}v_j+\sum_{j\in I_2}v_j=\sum_{j\in I_1}u_i+\sum_{j\in I_2}w_j=v$ proving that $\sum_{j\in I_1}V_j+\sum_{j\in I_2}V_j\subseteq\sum_{j\in I}V_j$ □

Theorem 10.363. *If V is a vector space over a field F and $\{V_i\}_{i\in I}$ is a **finite** family of non empty sub spaces then $\sum_{i\in I} V_i$ is a non empty subspace of V and thus a vector space.*

Proof. First as $0=\sum_{i\in I} 0$ we have that $0\in\sum_{i\in I} V_i$ and thus $\sum_{i\in I} V_i\neq\emptyset$. Second if $\alpha,\beta\in F, x,y\in\sum_{i\in I} V_i$ then there exists $\{x_i\}_{i\in I},\{y_i\}_{i\in I}$ such that $x=\sum_{i\in I} x_i$, $y=\sum_{i\in I} y_i$ and thus we have that

$$\begin{aligned}\alpha\cdot x+\beta\cdot y &= \alpha\cdot\left(\sum_{i\in I} x_i\right)+\beta\cdot\left(\sum_{i\in I} y_i\right)\\ &= \left(\sum_{i\in I}(\alpha\cdot x_i)\right)+\left(\sum_{i\in I}(\beta\cdot y_i)\right)\\ &= \sum_{i\in I}(\alpha\cdot x_i+\beta\cdot y_i)\\ &\in \sum_{i\in I} V_i\,(\text{as}\,\alpha\,x_i+\beta\cdot y_i\in V_i)\end{aligned}$$

□

Definition 10.364. *If V is a vector space over a field F then $V=\sum_{i\in I}^{\oplus} V_i$ if $\{V_i\}_{i\in I}$ is a **finite** family of non empty sub spaces and $\forall v\in V$ there exists a **unique** family $\{v_i\}_{i\in I}$ in V such that $\forall i\in I$ we have $v_i\in V_i$ and $v=\sum_{i\in I} v_i$ (note that the sum exists because I is finite).*

If $I=\{1,\ldots,n\}$ then we sometimes write $V=V_1\oplus\ldots.\oplus V_n$ instead of $V=\sum_{i\in I}^{\oplus} V_i$ and in the particular case of $I=\{1,\ldots,2\}$ as $V=V_1\oplus V_2$

Example 10.365. If V is a vector space over a field V then $\{V_i\}_{i\in\{1\}}$ where $V_1=V$ is a finite family of non empty sub spaces such that $\forall v\in V$ the unique $\{v_i\}_{i\in\{1\}}$ where $v=v_1$ has $v=v_1=\sum_{i\in\{1\}} v_i$ so that $V=\sum_{i\in\{1\}}^{\oplus} V_i$

Theorem 10.366. *If V is a vector space over a field F and $\{V_i\}_{i\in I}$ is a finite family of non empty sub spaces of V then $V=\sum_{i\in I}^{\oplus} V_i$ if and only if $V=\sum_{i\in I} V_i$ and $\forall i\in I$ we have that $V_i\bigcap\left(\sum_{j\in I\setminus\{i\}} V_j\right)=\{0\}$ where 0 is the neutral element of V*

Proof.

1. **(⇒)** As $V=\sum_{i\in I}^{\oplus} V_i$ we have by definition that $V\subseteq\sum_{i\in I} V_i\subseteq V\Rightarrow V=\sum_{i\in I} V_i$. Let $i\in I$ and take $v\in V_i\bigcap\left(\sum_{j\in I\setminus\{i\}} V_j\right)$ then $v\in V_i$ and there exists a $\{v_j\}_{j\in I\setminus\{i\}}$ such that $\forall j\in I\setminus\{i\}$ we have $v_j\in V_j$ and $v=\sum_{j\in I\setminus\{i\}} v_j$. Define now $\{w_j\}_{j\in I}$ where $w_j=\begin{cases} v_j \text{ if } j\in I\setminus\{i\}\\ -v \text{ if } j=i\end{cases}$ then we have that $\sum_{j\in I} w_i=\left(\sum_{j\in I\setminus\{i\}} w_j\right)+\left(\sum_{j\in\{i\}} w_j\right)=\left(\sum_{j\in I\setminus\{j\}} v_j\right)+(-v)=v+(-v)=0$ then as $0=\sum_{j\in I} 0$ we have by the uniqueness of the expansion of 0 that $\forall j\in I$ we have that $w_j=0$ and thus that $v=-w_i=-0=0$.

2. **($\Leftarrow$)** As $V=\sum_{i\in I} V_i$ we have if $v\in V$ that there exists a $\{v_i\}_{i\in I}$ with $\forall i\in I\vDash v_i\in V_i$ and $v=\sum_{i\in I} v_i$. Suppose now that $\{w_i\}_{i\in I}$ is another family with $\forall i\in I\vDash w_i\in V_i$ and $v=\sum_{i\in I} w_i$ which is different from $\{v_i\}_{i\in I}$ so that $\exists i_0\in I$ such that $v_{i_0}\neq w_{i_0}$. then we have that $0=v+(-v)=(\sum_{i\in I} v_i)+(-\sum_{i\in I} w_i)=\sum_{i\in I}(v_i-w_i)=\left(\sum_{i\in I\setminus\{i_0\}}(v_i-w_i)\right)+\left(\sum_{i\in\{i_0\}}(v_i-w_i)\right)=\left(\sum_{i\in I\setminus\{i_0\}}(v_i-w_i)\right)+(v_{i_0}-w_{i_0})\Rightarrow w_{i_0}-v_{i_0}=\sum_{i\in I\setminus\{i_0\}}(v_i-w_i)\in\sum_{i\in I\setminus\{i_0\}}(v_i-w_i)\Rightarrow w_{i_0}-v_{i_0}\in V_{i_0}\bigcap\left(\sum_{i\in I\setminus\{i_0\}} V_i\right)$ which by the assumption means that $w_{i_0}-v_{i_0}=0$ contradicting $w_{i_0}\neq v_{i_0}$. This proves then by contradiction that $\forall i\in I$ we have $v_i=w_i$ proving uniqueness of the expansion. □

Theorem 10.367. *If V is a vector space over a field, $\{V_i\}_{i\in I}$ a family of sub spaces such that $V=\sum_{i\in I}^{\oplus} V_{h(i)}$ where $h: J\to I$ is a bijection then $\sum_{i\in I}^{\oplus} V_{h(i)}=\sum_{j\in I}^{\oplus} V_i$*

Proof. By the definition of direct sum we must have $V=\sum_{i\in I} V_{h(i)}$ and then by 10.360 we have $V=\sum_{j\in J} V_j$. If now $j\in J$ and $V_j\bigcap\sum_{i\in J\setminus\{j\}} V_i$. Define $i_j=h^{-1}(j)\Rightarrow h(i_j)=j$ and take then $f=h_{|I\setminus\{i_j\}}: I\setminus\{i_j\}\to J\setminus\{j\}$ then f is a bijection and we have $\sum_{i\in I\setminus\{i_j\}} V_{h(i)}=\sum_{j\in I\setminus\{i_j\}} V_{f(i)} \underset{10.360}{=} \sum_{j\in J\setminus\{j\}} V_j$ so we have that $V_j\bigcap\sum_{i\in J\setminus\{j\}} V_i=V_{h(i_j)}\bigcap\sum_{i\in I\setminus\{i_j\}} V_{h(i)} \underset{V=\sum_{i\in I}^{\oplus} V_{h(i)}}{=} \{0\}$ □

Theorem 10.368. *If V is a vector space over F and $\{V_i\}_{i\in I}$ is a finite family of non empty sub spaces of V then (as by 10.363 $\sum_{j\in I\setminus\{i\}} V_j$ is a vector space over F (with $\{V_j\}_{j\in I\setminus\{i\}}$ as non empty subsets)) $\sum_{j\in I\setminus\{i\}} V_i=\sum_{j\in I\setminus\{i\}}^{\oplus} V_j$ and $V=\left(\sum_{j\in I\setminus\{i\}}^{\oplus} V_j\right)\oplus V_i$*

Proof.

1. **($\sum_{j\in I\setminus\{i\}} V_i=\sum_{j\in I\setminus\{i\}}^{\oplus} V_i$)** As we trivially have that $\sum_{j\in I\setminus\{i\}} V_i=\sum_{j\in I\setminus\{i\}} V_i$ we must prove that $\forall k\in I\setminus\{i\}$ we have that $V_k\bigcap\left(\sum_{j\in(I\setminus\{i\})\setminus\{k\}} V_j\right)=\{0\}$. Now if $x\in V_k\bigcap\left(\sum_{j\in(I\setminus\{i\})\setminus\{k\}} V_j\right)$ then $x\in V_k$ and there exists a $\{x_i\}_{j\in(I\setminus\{i\})\setminus\{k\}}$ such that $\forall j\in(I\setminus\{i\})\setminus\{k\}$ we have $x_j\in V_j$ and $x=\sum_{j\in(I\setminus\{i\})\setminus\{k\}} x_j$, if we define then $\{y_j\}_{j\in I\setminus\{k\}}$ by $y_j=\begin{cases} x_j \text{ if } j\neq\{i\} \\ 0 \text{ if } j=i\end{cases}\in V_j$ then $x=\sum_{j\in(I\setminus\{i\})\setminus\{k\}} x_j+0=\sum_{j\in(I\setminus\{i\})\setminus\{k\}} y_j+y_i=\sum_{j\in I\setminus\{k\}} y_j\in\sum_{I\setminus\{k\}} V_j$ giving that $x\in V_k\bigcap\left(\sum_{j\in I\setminus\{k\}} V_j\right)=\{0\}$ so that $x=0$.

2. **($V=\left(\sum_{j\in I\setminus\{i\}}^{\oplus} V_j\right)\oplus V_i$)** Using 10.361 we have that $\left(\sum_{k\in I\setminus\{i\}} V_j\right)+V_i$ so we must prove that if $\left(\sum_{j\in I\setminus\{i\}} V_j\right)\bigcap V_i=\{0\}$ which is true because $V=\sum_{i\in I}^{\oplus} V_i$ by 10.366. □

Lemma 10.369. *If V is a vector space over a field F, $V_1,V_2\subseteq V$ non empty finite dimensional sub spaces of V (which may be trivial) then* $\dim(V_1\oplus V_2)=\dim(V_1)+\dim(V_2)$

Proof. We have the following cases for V_1, V_2 to consider:

V_1, V_2 are trivial. If V_1, V_2 are trivial then $V_1 = \{0\} = V_2$ hence $V_1 \oplus V_2 = \{0\}$ so that $\dim(V \oplus V_2) = 0 = 0 + 0 + \dim(V_1) + \dim(V_2)$

V_1 is trivial, V_2 is not trivial. Then $V_1 = \{0\}$ and V_2 has a basis $\{e_i\}_{i\in\{1,\ldots,n\}}$. If $x \in V_1 \oplus V_2$ there exists a $x_1 \in V_1, x_2 \in V_2$ such that $x = x_1 + x_2 \underset{V_1=\{0\}}{=} 0 + x_2$ proving that $x \in V_2$ hence there exists a $\{\alpha_i\}_{i\in\{1,\ldots,n\}}$ such that $x = \sum_{i=1}^n \alpha_i \cdot e_i$. As $\{e_i\}_{i\in\{1,\ldots,n\}}$ are by definition linear independent we have that $\{e_i\}_{i\in\{1,\ldots,n\}}$ is a basis of $V_1 \oplus V_2$. So $\dim(V_1 \oplus V_2) = n = \dim(V_2) = 0 + \dim(V_2) = \dim(V_1) + \dim(V_2)$.

V_1 is not trivial, V_2 is trivial. This is the same as the above if we interchage the roles of V_1 and V_2

V_1, V_2 are not trivial. Define then $\varphi: \prod_{i\in\{1,\ldots,2\}} V_i \rightarrow V_1 \oplus V_2$ by $\varphi(v_1, v_2) = v_1 + v_2$ then φ is a isomorphism:

injectivity. If $\varphi(v_1, v_2) = \varphi(u_1, u_2) \Rightarrow v_1 + v_2 = u_1 + u_2 \underset{\text{uniquenessofsum}}{\Rightarrow} (v_1, v_2) = (u_1, u_2)$

surjectivity. If $v \in V_1 \circ V_2$ then there exists a $v_1 \in V_1, v_2 \in V_2$ such that $v = v_1 + v_2 = \varphi(v_1, v_2)$

linearity. If $(v_1, v_2), (u_1, u_2) \in \prod_{i\in\{1,\ldots,2\}} V_i$ and $\alpha, \beta \in F$ then $\varphi(\alpha \cdot (v_1, v_2) + \beta \cdot (u_1, u_2)) = \varphi(\alpha \cdot v_1 + \beta \cdot u_1, \alpha \cdot v_2 + \beta \cdot u_2) = \alpha \cdot v_1 + \beta \cdot u_1 + \alpha \cdot v_2 + \beta \cdot u_2 = \alpha \cdot (v_1 + u_1) + \beta \cdot (v_2 + u_2) = \alpha \cdot \varphi(v_1, u_1) + \beta \cdot (v_2, u_2)$

Hence $\dim(V_1 \oplus V_2) \underset{10.220 \text{ and } \varphi \text{ is a isomorphisme}}{=} \dim\left(\prod_{i\in\{1,\ldots,2\}} V_i\right) \underset{10.202}{=} \sum_{i=1}^2 \dim(V_i) = \dim(V_1) + \dim(V_2)$ □

Theorem 10.370. *If V is a vector spaces over a field F, $n \in \mathbb{N} \setminus \{1\}$, and $\{V_i\}_{i\in\{1,\ldots,n\}}$ a finite family of non empty finite dimensional such that $V = V_1 \oplus \cdots \oplus V_n$ then* $\dim(V) = \sum_{i=1}^n \dim(V_i)$

Proof. We prove this by induction on n so let $\mathcal{S} = \{n \in \{2, \ldots, \infty\} |$if V is *a* vector space $\{V_i\}_{i\in\{1,\ldots,n\}}$ finite dimensional non empty subspace of V such that $V = V_1 \oplus \cdots \oplus V_n$ then $\dim(V) = \sum_{i=1}^n \dim(V_i)$ $\}$ then we have

$2 \in \mathcal{S}$. This follows from the previous lemma (see 10.369)

$n \in S \Rightarrow n + 1 \in \mathcal{S}$. If $\{V_i\}_{i\in\{1,\ldots,n+1\}}$ is a family of non empty finite subspaces of V then we have by 10.368 that $V = (V_1 \oplus \cdots \oplus V_{n+1}) = (V_1 \oplus \cdots \oplus V_n) \oplus V_{n+1}$ so that $\dim(V) = \dim((V_1 \oplus \cdots \oplus V_n) \oplus V_{n+1}) \underset{10.369}{=} \dim(V_1 \oplus \cdots \oplus V_n) + \dim(V_{n+1}) \underset{n\in\mathcal{S}}{=} (\sum_{i=1}^n \dim(V_i)) + \dim(V_{n+1}) = \sum_{i=1}^{n+1} \dim(V_i)$ proving that $n + 1 \in \mathcal{S}$

Induction finishes then the proof. □

This theorem can be generalized also

Theorem 10.371. *If V is a vector space over a field F and $\{V_i\}_{i\in I}$ is a finite family of non empty sub spaces of V and $I = I_1 \bigcup I_2, I_1 \bigcap I_2 = \emptyset$ (by 10.363 we have then that $\sum_{i\in I_1} V_i$, $\sum_{i\in I_2} V_i$ are vector spaces over F with sub spaces $\{V_i\}_{i\in I_1}, \{V_i\}_{i\in I_2}$). Then $\sum_{i\in I_1} V_i = \sum_{i\in I_1}^{\oplus} V_i$, $\sum_{i\in I_2} V_i = \sum_{i\in I_2}^{\oplus} V_i$ (more general if $J \subseteq I$ then $\sum_{j\in J}^{\oplus} V_j$ exist) and $V = (\sum_{i\in I_1}^{\oplus} V_i) \oplus (\sum_{i\in I_2}^{\oplus} V_i)$*

Proof.

1. Let $J \subseteq I$ and prove that $\sum_{j\in J} V_j = \sum_{j\in J}^{\oplus} V_j$ then as we trivially have $\sum_{j\in J} V_j = \sum_{j\in J} V_j$ we must prove that $\forall i \in J$ we have $V_i \bigcap \left(\sum_{j\in J\setminus\{i\}} V_j\right) = \{0\}$. Now if $x \in V_i \bigcap \left(\sum_{j\in J\setminus\{i\}} V_j\right)$ then $x \in V_i$ and $x = \sum_{j\in J\setminus\{i\}} x_j$ where for the family $\{x_j\}_{j\in J\setminus\{i\}}$ we have $x_j \in X_j$. Define now $\{y_j\}_{j\in I\setminus\{i\}}$ by $y_j = \begin{cases} x_j \text{ if } j \in J\setminus\{i\} \\ 0 \text{ if } j \in (I\setminus\{i\})\setminus(J\setminus\{i\}) \end{cases}$ then $x = \sum_{j\in J\setminus\{i\}} x_j + 0 = \sum_{j\in J\setminus\{i\}} x_j + \sum_{j\in(I\setminus\{i\})\setminus(J\setminus\{i\})} 0 = \sum_{j\in J\setminus\{i\}} y_i + \sum_{j\in(I\setminus\{i\})\setminus(J\setminus\{i\})} y_j = \sum_{j\in I\setminus\{i\}} y_j \in \sum_{j\in I\setminus\{i\}} V_i$ giving that $x \in V_i \bigcap \left(\sum_{j\in I\setminus\{i\}} V_j\right) = \{0\}$ giving that $x = 0$ or $V_i \bigcap \left(\sum_{j\in I\setminus\{i\}} V_j\right)$. Applying the above for $J = I_1, I_2$ proves the first assertion of the theorem.

2. Using 10.362 we have that $V = \sum_{j\in I}^{\oplus} V_j \underset{\text{definition of direct sum}}{=} \sum_{j\in I} V_j = \sum_{j\in I_1} V_j + \sum_{j\in I_2} V_j$. So we must only prove that $(\sum_{j\in I_1} V_j) \bigcap (\sum_{j\in I_2} V_j) = \{0\}$ and do this using proof by contradiction. If $x \in (\sum_{j\in I_1} V_j) \bigcap (\sum_{j\in I_2} V_j)$ and $x \neq 0$ then $0 \neq x = \sum_{i\in I_1} x_i = \sum_{i\in I_2} y_i$ where $\forall i \in I_1 \vDash x_i \in V_i$, $\forall i \in I_2 \vDash x_i \in V_i$. As $x \neq 0$ there exists a $i \in I_1$ such that $V_i \ni x_i \neq 0$ define now $\{v_j\}_{j\in I\setminus\{i\}}$ by $v_j = \begin{cases} -x_j \text{ if } j \in I_1\setminus\{i\} \\ y_j \text{ if } j \in I_2 \end{cases}$ then $\sum_{j\in I\setminus\{i\}} V_j \ni \sum_{j\in I\setminus\{i\}} v_j = \sum_{j\in I_1\setminus\{i\}} v_j + \sum_{j\in I_2} v_j = \sum_{j\in I_1\setminus\{i\}} (-x_i) + \sum_{j\in I_2} y_j = -\sum_{j\in I_1\setminus\{i\}} x_j + x = -\left(\sum_{j\in I_1\setminus\{i\}} x_i + x_i\right) + x_i + x = -x + x_i + x = x_i$ proving that $0 \neq x_i \in V_i \bigcap \left(\sum_{j\in I\setminus\{i\}} V_j\right) = \{0\}$ which is a contradiction. □

Theorem 10.372. *If V is a vector space over a field F and $\{V_i\}_{i\in I}$ is a finite family of non empty non trivial sub spaces of V and $I = \bigcup_{i\in\{1,\ldots,m\}} I_i$ where I_i's are pairwise distinct then we have $\sum_{i\in\{1,\ldots,m\}} (\sum_{j\in I_i}^{\oplus} V_j) = \sum_{i\in\{1,\ldots,m\}}^{\oplus} (\sum_{j\in I_i}^{\oplus} V_j) = \sum_{i\in I}^{\otimes} V_i$ (if $\sum_{i\in I}^{\otimes} V_i$ exist)*

Proof. The proof is by induction, let $S = \left\{m \in \mathbb{N} \middle| \text{If } I = \bigcup_{i<\{1,\ldots,m\}} I_i \text{ (pairwise disjoint union) and } \sum_{i\in I}^{\oplus} V_i \text{ exist then } \sum_{i\in I}^{\oplus} V_i = \sum_{i\in\{1,\ldots,m\}} \left(\sum_{j\in I_i^{\oplus}} V_j\right)\right\}$ then we have

1. $(\boldsymbol{m = 1})$ Here $I = I_1$ and $\sum_{i\in\{1,\ldots,1\}} (\sum_{j\in I_i}^{\oplus} V_j) = \sum_{j\in I_1}^{\oplus} V_j = \sum_{j\in I}^{\oplus} V_j \Rightarrow 1 \in S$

2. **($m \in S$)** If $I = \bigcup_{j \in \{1, \ldots, m+1\}} I_j$ (pairwise disjoint union) and $\sum_{j \in I}^{\oplus} V_j$ exist take then $J = \bigcup_{i \in \{1, \ldots, m\}} I_i$ then by the previous theorem $\sum_{j \in J}^{\oplus} V_j$ exists and as $I = J \bigcup I_{m+1}$ and trivially $J \bigcap I_{m+1} = \emptyset$ we have

$$\begin{aligned}\sum_{i \in I}^{\oplus} V_i &= \left(\sum_{j \in J}^{\oplus} V_j\right) \oplus \left(\sum_{i \in I_{m+1}}^{\oplus} V_i\right) \\ &\underset{m \in S}{=} \left(\sum_{i \in \{1, \ldots, m\}}^{\oplus} \left(\sum_{j \in I_i}^{\oplus} V_i\right)\right) \oplus \left(\sum_{j \in I_{m+1}}^{\oplus} V_i\right) \\ &\underset{10.368}{=} \sum_{i \in \{1, \ldots, m+1\}}^{\oplus} \left(\sum_{j \in I_i}^{\oplus} V_i\right)\end{aligned}$$

proving that $m + 1 \in S$

By induction we have then that $S = \mathbb{N}$ which proves the theorem.

□

10.12 Tensor product of vector spaces

Definition 10.373. (Tensor Product) *If $\{V_i\}_{i \in \{1, \ldots, n\}}$ is a finite non empty family of vector spaces over a field F then a* ***tensor product*** *of $\{V_i\}_{i \in \{1, \ldots, n\}}$ is a pair $\langle P, \nu \rangle$ where*

1. *P is a vector space over the field F*
2. *$\nu \in \mathrm{Hom}(V_1, \ldots, V_n; P)$*
3. *$\mathcal{S}\left(\nu\left(\prod_{i \in \{1, \ldots, n\}} V_i\right)\right) = P$*
4. *If $\varphi \in \mathrm{Hom}(V_1, \ldots, V_n; U)$ then there exists a $h \in \mathrm{Hom}(P, U)$ such that $\varphi = h \circ \nu$*

We prove now that the tensor product is uniquely determined to within a canonical isomorphism.

Theorem 10.374. *If $\{V_i\}_{i \in \{1, \ldots, n\}}$ is a finite non empty family of vector spaces over a field F and $\langle P, \nu \rangle$ and $\langle Q, \mu \rangle$ be tensor products of $\{V_i\}_{i \in \{1, \ldots, n\}}$ then there exists a* ***unique*** *bijective linear function (= linear isomorphism) $k: P \rightarrow Q$ such that $k \circ \nu = \mu$*

Proof. As $\mu \in \mathrm{Hom}(V_1, \ldots, V_n; Q)$ we have by the fact that $\langle P, \nu \rangle$ is a tensor product there exists a $k \in \mathrm{Hom}(P, Q)$ such that

$$k \circ \nu = \mu. \tag{10.136}$$

Also as $\nu \in \mathrm{Hom}(V_1, \ldots, V_n; P)$ we have by the fact that $\langle Q, \mu\rangle$ is a tensor product there exist a $h \in \mathrm{Hom}(Q, P)$ such that

$$h \circ \mu = \nu \tag{10.137}$$

Using 10.136 and 10.137 we have that

$$k \circ h \circ \mu = \mu \tag{10.138}$$

$$h \circ k \circ \nu = \nu \tag{10.139}$$

If now $x \in P$ then as $\mathcal{S}\left(\nu\left(\prod_{i \in \{1, \ldots, n\}} V_i\right)\right) = P$ there exists a $\{\alpha_i\}_{i \in \{1, \ldots, k\}}$ in F and a $\{y_i\}_{i \in \{1, \ldots, k\}}$ in $\nu\left(\prod_{i \in \{1, \ldots, k\}} V_i\right)$ such that $x = \sum_{i \in \{1, \ldots, k\}} \alpha_i \cdot y_i$, now $\forall i \in \{1, \ldots, k\}$ there exists a unique x_i such that $\nu(x_i) = y_i$, which defines a family $\{x_i\}_{i \in \{1, \ldots, k\}}$ in $\prod_{i \in \{1, \ldots, n\}} V_i$ such that $x = \sum_{i \in \{1, \ldots, k\}} \alpha_i \cdot \nu(x_i)$ so that $(h \circ k)(x) = (h \circ k)\left(\sum_{i \in \{1, \ldots, k\}} \alpha_i \cdot \nu(x_i)\right) \underset{h \circ k \text{ is linear}}{=} \sum_{i \in \{1, \ldots, k\}} \alpha_i \cdot (h \circ k)(\nu(x_i)) = \sum_{i \in \{1, \ldots, k\}} \alpha_i \cdot (h \circ k \circ \nu)(x_i) \underset{10.139}{=} \sum_{i \in \{1, \ldots, k\}} \alpha_i \cdot \nu(x_i) = x$ proving that

$$h \circ k = 1_P \tag{10.140}$$

If now $x \in Q$ then as $\mathcal{S}\left(\mu\left(\prod_{i \in \{1, \ldots, n\}} V_i\right)\right) = Q$ there exists a $\{\alpha_i\}_{i \in \{1, \ldots, k\}}$ in F and a $\{y_i\}_{i \in \{1, \ldots, k\}}$ in $\mu\left(\prod_{i \in \{1, \ldots, k\}} V_i\right)$ such that $x = \sum_{i \in \{1, \ldots, k\}} \alpha_i \cdot y_i$, now $\forall i \in \{1, \ldots, k\}$ there exists a unique x_i such that $\mu(x_i) = y_i$, which defines a family $\{x_i\}_{i \in \{1, \ldots, k\}}$ in $\prod_{i \in \{1, \ldots, n\}} V_i$ such that $x = \sum_{i \in \{1, \ldots, k\}} \alpha_i \cdot \mu(x_i)$ so that $(k \circ h)(x) = (k \circ h)\left(\sum_{i \in \{1, \ldots, k\}} \alpha_i \cdot \mu(x_i)\right) \underset{h \circ k \text{ is linear}}{=} \sum_{i \in \{1, \ldots, k\}} \alpha_i \cdot (k \circ h)(\mu(x_i)) = \sum_{i \in \{1, \ldots, k\}} \alpha_i \cdot (k \circ h \circ \mu)(x_i) \underset{10.139}{=} \sum_{i \in \{1, \ldots, k\}} \alpha_i \cdot \mu(x_i) = x$ proving that

$$k \circ h = 1_Q \tag{10.141}$$

This proves that k and h are bijections (and as they are linear) that k, h are linear isomorphism's.

To prove uniqueness, suppose that there is another bijection $k' \in \mathrm{Hom}(P, Q)$ such that $k' \circ \nu = \mu$ then $k' \circ \nu = k \circ \nu$, again as $x \in P = \mathcal{S}\left(\nu\left(\prod_{i \in \{1, \ldots, n\}} V_i\right)\right)$ there exists a $\{\alpha_i\}_{i \in \{1, \ldots, n\}}$ in F and a $\{x_i\}_{i \in \{1, \ldots, n\}}$ in $\prod_{i \in \{1, \ldots, n\}} V_i$ such that $x = \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot \nu(x_i)$ so that $k'(x) = k'\left(\sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot \nu(x_i)\right) = \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot k'(\nu(x_i)) \underset{k' \circ \nu = k \circ \nu}{=} \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot k(\nu(x_i)) = k\left(\sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot \nu(x_i)\right) = k(x)$ so that we have $k' = k$ $\square$

Because of the previous theorem it makes sense to introduce a special notation for the tensor product.

Notation 10.375. *If $\langle P, \nu\rangle$ is a tensor product of $\{V_i\}_{i \in \{1, \ldots, n\}}$ then we note P as $V_1 \otimes \ldots \otimes V_n$ or $\otimes_{i \in \{1, \ldots, n\}} V_i$ and ν as $\otimes$. Also if $(x_1, \ldots, x_n) \in \prod_{i \in \{1, \ldots, n\}} V_i$ then $\nu(x)$ is noted as $x_1 \otimes \ldots \otimes x_n$.*

Example 10.376. If $\{V_i\}_{i\in\{1\}}$ is a finite non empty family of vector spaces over a field F then $\langle V_1, 1_{V_1}\rangle$ forms a tensor product.

Proof.

1. V_1 is trivially a vector space over F
2. $1_{V_1}\in \mathrm{Hom}(V_1;V_1)$ trivially
3. $\mathcal{S}\left(1_{V_1}\left(\prod_{i\in\{1\}} V_i\right)\right)=\mathcal{S}(1_{V_1}(V_1))=\mathcal{S}(V_1)=V_1$
4. If $\varphi\in\mathrm{Hom}(V_1,U)$ then $\varphi\in\mathrm{Hom}(V_1,U)$ satisfies $\varphi=\varphi\circ 1_{V_1}$ □

Next we prove a theorem that specifies the zero elements in tensor product considered as a vector space

Theorem 10.377. *If $\{V_i\}_{i\in\{1,\ldots,m\}}$ is a finite non empty family of vector spaces over a field F with multiplicity zero then if $(v_1,\ldots,v_m)\in\prod_{i\in\{1,\ldots,\ldots m\}} V_i$ we have $(v_1\otimes\ldots\otimes v_m)=0$ if and only if $\exists i\in\{1,\ldots,m\}$ with $v_i=0$*

Proof.

($\Rightarrow$) This is proved by contradiction, so let $v_1\otimes\ldots\otimes v_m=0$ with $\forall i\in\{1,\ldots,m\}$ we have $v_i\neq 0$ then by 10.176 there exists linear mappings $\varphi_i\colon V_i\rightarrow F$ with $\varphi_i(v_i)\neq 0$. This defines a multilinear map $\varphi\colon\prod_{i\in\{1,\ldots,m\}} V_i\rightarrow F$ defined by $\varphi(v_1,\ldots,v_m)=\prod_{i\in\{1,\ldots,m\}}\varphi_i(v_i)$, by the definition of a tensor space there exist then a linear mapping $h\colon V_1\otimes\ldots\otimes V_m\rightarrow F$ such that $f\circ\otimes=\varphi$. So $f(v_1\otimes\ldots\otimes v_m)=\varphi(v_1,\ldots,v_m)=1\underset{v_1\otimes\ldots\otimes v_m=0\,\mathrm{and}\,h\,\mathrm{is\,linear}}{\Rightarrow}0=1$ a contradiction as F has multiplicity of zero. So we must conclude that $\exists i\in\{1,\ldots,m\}$ such that $v_i=0$

($\Leftarrow$) If $\exists i\in\{1,\ldots,m\}$ with $v_i=0$ then $v_i=0\cdot v_i$ so that $v_1\otimes\ldots.\otimes v_m=v_1\otimes\ldots\otimes v_{i-1}\otimes(0\cdot v_i)\otimes v_{i+1},\ldots,\otimes v_m=0\cdot(v_1\otimes\ldots\otimes v_m)=0$ □

Theorem 10.378. *If $\{V_i\}_{i\in\{1,\ldots,n\}}$ is a finite non empty family of vector spaces over a field F and $V_1\otimes\ldots\otimes V_n$ a tensor product then $\forall x\in V_1\otimes\ldots\otimes V_n$ there exist a $\{x_{1,i}\otimes\ldots\otimes x_{n,i}\}_{i\in k}$ such that $x=\sum_{i\in\{1,\ldots,k\}}(x_{1,i}\otimes\ldots\otimes x_{n,i})$ where k is minimal.*

Proof. As $V_1\otimes\ldots\otimes V_n=\mathcal{S}\left(\otimes\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)\right)$ we have that for every $x\in V_1\otimes\ldots\otimes V_n$ there exists a $\{\lambda_i\}_{i\in\{1,\ldots,k\}}$ and $\{y_{1,i}\otimes\ldots\otimes y_{n,i}\}_{i\in\{1,\ldots,k\}}$ such that $x=\sum_{i\in\{1,\ldots,k\}}\lambda_i\cdot(y_{1,i}\otimes\ldots\otimes y_{n,i})$ then using the multilinearity of $\otimes$ and the fact that V_1 is a vector space we have that $\lambda_i\cdot(y_{1,i}\otimes\ldots\otimes y_{n,i})=(\lambda_i\cdot y_{1,i})\otimes\ldots\otimes y_{n,i}\in\otimes(\Pi_{i\in\{1,\ldots,n\}}V_i)$ so if $x_{1,i}\otimes\ldots\otimes x_{n,i}=(\lambda_i\cdot y_{1,i})\otimes\ldots\otimes y_{n,i}\in\otimes(\Pi_{i\in\{1,\ldots,n\}}V_i)$ we have then that $x=\sum_{i\in\{1,\ldots,k\}}x_{1,i}\otimes\ldots\otimes x_{n,i}$. So there exists a $\{x_{1,i}\otimes\ldots\otimes x_{n,i}\}_{i\in\{1,\ldots,k\}}$ such that $x=\sum_{i\in\{1,\ldots,k\}}x_{1,i}\otimes\ldots\otimes x_{n,i}$ proving that $N=\left\{m\in\mathbb{N}|\text{there exists } a\ \{x_{1,i}\otimes\ldots\otimes x_{n,i}\}_{i\in\{1,\ldots,m\}}\text{ such that } x=\sum_{i\in\{1,\ldots,m\}}x_{1,i}\otimes\ldots\otimes x_{n,i}\right\}\neq\emptyset$ so that $k=\min(N)$ (see 4.52) exists so that there exists a $\{x_{1,i}\otimes\ldots\otimes x_{n,i}\}_{i\in\{1,\ldots,k\}}$ so that $x=\sum_{i\in\{1,\ldots,k\}}(x_{1,i}\otimes\ldots\otimes x_{n,i})$ proving the theorem. □

We proceed now to prove that there exists always a tensor product.

Remember that given a set X and a field $\langle F, +, \cdot\rangle$ we can using 10.169 create a vector space $\langle \mathcal{F}(X, F), +, \cdot\rangle$ over the field $\langle F, +, \cdot\rangle$ called the **free vector space over X** such that $\{\delta_a\}_{a\in X}$ forms a basis of $\langle \mathcal{F}(X, F), +, \cdot\rangle$. Or if F has characterization 0 we have that X is embedded in $\mathcal{F}(X, F)$ by the injection δ and forms the basis of $\mathcal{F}(X, F)$.

Given a finite family of vector spaces $\{V_i\}_{i\in\{1,\ldots,n\}}$ over a field F we define now the tensor product $V_1\otimes\ldots\otimes V_n$ over the field F using the following steps.

First we form the free vector space over $\prod_{i\in\{1,\ldots,n\}} V_i$ (We consider this as just a set with no additional structure).

Definition 10.379. *If $\{V_i\}_{i\in\{1,\ldots,n\}}$ is a finite non empty family of vector spaces over a field F define then $\mathcal{N}\left(\prod_{i\in\{1,\ldots,n\}} V_i, F\right)\subseteq \mathcal{F}\left(\prod_{i\in\{1,\ldots,n\}} V_i, F\right)$ by $\mathcal{N}\left(\prod_{i\in\{1,\ldots,n\}} V_i, F\right) = \bigcup_{i\in\{1,\ldots,n\}} \left\{\delta_{(x_1,\ldots,x_{i-1},\alpha\cdot x+\beta\cdot y)} - \alpha\cdot \delta_{(x_1,\ldots,x_{i-1},x,x_{i+1},\ldots,x_n)} - \beta\cdot\delta_{(x_1,\ldots,x_{i-1},y,x_{i+1},\ldots,x_n)} | (x_1,\ldots,x_{i-1},\alpha\cdot x+\beta\cdot y, x_{i+1})\in \prod_{i\in\{1,\ldots,n\}} V_i\wedge\alpha,\beta\in F\right\}$*

We are now ready to define $V_1\otimes\ldots\otimes V_n$

Definition 10.380. *If $\{V_i\}_{i\in\{1,\ldots,n\}}$ is a finite family of vector spaces over a field F take then the free vector space $\left\langle\mathcal{F}\left(\prod_{i\in\{1,\ldots,n\}} V_i, +,\cdot\right)\right\rangle$ over the field F and the sub vector space $\mathcal{S}\left(\mathcal{N}\left(\prod_{i\in\{1,\ldots,n\}} V_i, F\right)\right)$ then $\langle V_1\otimes\ldots\otimes V_n, +,\cdot\rangle$ is defined to by the factor space (see 10.191) $\left\langle\mathcal{F}\left(\prod_{i\in\{1,\ldots,n\}} V_i, +,\cdot\right)/\mathcal{S}\left(\mathcal{N}\left(\prod_{i\in\{1,\ldots,n\}} V_i, F\right)\right), +, \cdot\right\rangle$.*

Definition 10.381. *If $\{V_i\}_{i\in\{1,\ldots,n\}}$ is a finite family of vector spaces over a field F then we define $\otimes: \prod_{i\in\{1,\ldots,n\}} \rightarrow V_1\otimes\ldots\otimes V_n$ by $(v_1,\ldots,v_n)\rightarrow v_1\otimes\ldots\otimes v_n = \pi(\delta(v_1,\ldots,v_n))$ so that $\otimes = \pi\circ\delta$ where we have defined*

1. *$\delta: \prod_{i\in\{1,\ldots,n\}} V_i\rightarrow\mathcal{F}\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)$ by $(x_1,\ldots,x_n)\rightarrow\delta_{(x_1,\ldots,x_n)}$ and $\delta_{(x_1,\ldots,x_n)}(r_1,\ldots,s) = \left\{\begin{array}{l} 1 \text{ if } (x_1,\ldots,x_n) = (r_1,\ldots,r_n)\\ 0 \text{ if } (x_1,\ldots,x_n)\neq(r_1,\ldots,r_n)\end{array}\right.$ (see 10.169)*
2. *$\pi: \mathcal{F}\left(\prod_{i\in\{1,\ldots,n\}} V_i, F\right)\rightarrow\mathcal{F}\left(\prod_{i\in\{1,\ldots,n\}} V_i, F\right)/\mathcal{S}\left(\mathcal{N}\left(\prod_{i\in\{1,\ldots,n\}} V_i, F\right)\right)$ is the canonical surjection $x\rightarrow\sim[x] = \left\{y | x-y\in\mathcal{S}\left(\mathcal{N}\left(\prod_{i\in\{1,\ldots,n\}} V_i, F\right)\right)\right\}$ which is linear by the definition of sum and scalar product on the quotient space (see 10.191)*

we call $v_1\otimes\ldots\otimes v_n$ the tensor product of the vectors $v_1,\ldots,v_n$

We prove now that the tensor product exist.

Theorem 10.382. *If $\{V_i\}_{i\in\{1,\ldots,n\}}$ is a non empty finite family of vector spaces over a field F then $\otimes: \prod_{i\in\{1,\ldots,n\}} V_i \rightarrow V_1 \otimes \ldots, \otimes V_n$ satisfies the axioms of a tensor product*

1. *$\otimes$ is multilinear or in other words $\otimes \in \mathrm{Hom}(V_1, \ldots, V_n; V_1 \otimes \ldots \otimes V_n)$ (in a simple notation this means $v_1 \otimes \ldots \otimes v_{i-1} \otimes \alpha \cdot x + \beta \cdot y \otimes v_{i+1} \otimes \ldots \otimes v_n = \alpha \cdot v_1 \otimes \ldots \otimes v_{i-1} \otimes x \otimes v_{i+1} \otimes \ldots \otimes v_n + \beta \cdot v_1 \otimes \ldots \otimes v_{i-1} \otimes y \otimes v_{i+1} \otimes \ldots \otimes v_n$*
2. *$\mathcal{S}\left(\otimes\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)\right) = V_1 \otimes \ldots \otimes V_n$*
3. **(Universal factorization property)** *If U is any vector space over F and $\varphi \in \mathrm{Hom}(V_1, \ldots, V_n; U)$ is arbitrary then there exist a $h \in \mathrm{Hom}(V_1 \otimes \ldots \otimes V_n, U)$ (a linear mapping) such that $\varphi = h \circ \otimes$*

Or in other words $\langle V_1 \otimes \ldots \otimes V_n, \otimes \rangle$ is a tensor product of $\{V_i)_{i\in\{1,\ldots,n\}}$.

Proof.

1. If $(v_1, \ldots, v_{i-1}, \alpha \cdot x + \beta \cdot y, v_{i+1}, \ldots, v_n) \in \prod_{i\in\{1,\ldots,n\}} V_i$ then $\delta(v_1, \ldots, v_{i-1}, \alpha \cdot x + \beta \cdot y, v_{i+1}, \ldots v_n) - \alpha \cdot \delta(v_1, \ldots, v_{i-1}, x, v_i, \ldots, v_n) - \beta \cdot \delta(v_1, \ldots, v_{i-1}, y, v_{i+1}, \ldots, v_n) \in \mathcal{N}\left(\prod_{i\in\{1,\ldots,n\}} V_i, F\right) \subseteq \mathcal{S}\left(\mathcal{N}\left(\prod_{i\in\{1,\ldots,n\}} V_i, F\right)\right)$ so that $0 = \pi(\delta(v_1, \ldots, v_{i-1}, \alpha \cdot x + \beta \cdot y, v_{i+1}, \ldots v_n) - \alpha \cdot \delta(v_1, \ldots, v_{i-1}, x, v_i, \ldots, v_n) - \beta \cdot \delta(v_1, \ldots, v_{i-1}, y, v_{i+1}, \ldots, v_n)) = \pi(\delta(v_1, \ldots, v_{i-1}, \alpha \cdot x + \beta \cdot y, v_{i+1}, \ldots v_n)) - \alpha \cdot \pi(\delta(v_1, \ldots, v_{i-1}, x, v_i, \ldots, v_n)) - \beta \cdot \pi(\delta(v_1, \ldots, v_{i-1}, y, v_{i+1}, \ldots, v_n)) \Rightarrow 0 = \otimes(v_1, \ldots, v_{i-1}, \alpha \cdot x + \beta \cdot y, v_{i+1}, \ldots, v_n) - \alpha \cdot \otimes(v_1, \ldots, v_{i-1}, x, v_{i+1}, \ldots, v_n) - \beta \cdot \otimes(v_1, \ldots, v_{i-1}, y, v_{i+1}, \ldots, v_n)$ giving that $\otimes(v_1, \ldots, v_{i-1}, \alpha \cdot x + \beta \cdot y, v_{i+1}, \ldots, v_n) = \alpha \cdot \otimes(v_1, \ldots, v_{i-1}, x, v_{i+1}, \ldots, v_n) + \beta \cdot \otimes(v_1, \ldots, v_{i-1}, y, v_{i+1}, \ldots, v_n)$ so that $\otimes \in \mathrm{Hom}(V_1, \ldots, V_n, V_1 \otimes \ldots \otimes V_n)$
2. As we already have that $\mathcal{S}\left(\otimes\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)\right) \subseteq V_1 \otimes \ldots \otimes V_n$ we must only prove that $V_1 \otimes \ldots \otimes V_n \subseteq \mathcal{S}\left(\otimes\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)\right)$. So let $x \in V_1 \otimes \ldots \otimes V_n$ then as π is a surjection there exists a $y \in \mathcal{F}\left(\prod_{i\in\{1,\ldots,n\}} V_i, F\right)$ such that $\pi(y) = x$. Now as $\delta\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)$ forms a basis of $\mathcal{F}\left(\prod_{i\in\{1,\ldots,n\}} V_i, F\right)$ (see 10.169) there exists a $\{\alpha_i\}_{i\in\{1,\ldots,n\}}$ and a $\left\{\delta\left(v_1^{(i)}, \ldots, v_n^{(i)}\right)\right\}_{i\in\{1,\ldots,n\}}$(see 10.132) such that $y = \sum_{i\in\{1,\ldots,n\}} \alpha_i \cdot \delta\left(v_1^{(i)}, \ldots, v_n^{(i)}\right)$ and thus $x = \pi\left(\sum_{i\in\{1,\ldots,n\}} \alpha_i \cdot \delta\left(v_1^{(i)}, \ldots, v_n^{(i)}\right)\right) = \sum_{i\in\{1,\ldots,n\}} \alpha_i \cdot \pi\left(\delta\left(v_1^{(i)}, \ldots, v_n^{(i)}\right)\right) = \sum_{i\in\{1,\ldots,n\}} \alpha_i \cdot v_1^{(i)} \otimes \ldots v_n^{(i)} \in \mathcal{S}\left(\otimes\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)\right)$
3. As $\left\{\delta_x | z \in \prod_{i\in\{1,\ldots,n\}} V_i\right\}$ is by 10.169 a basis for $\mathcal{F}\left(\prod_{i\in\{1,\ldots,n\}} V_i, F\right)$ we have if $x \in \mathcal{F}\left(\prod_{i\in\{1,\ldots,n\}} V_i, F\right)$ that there exists a unique $\left\{\alpha_z^{(x)}\right\}_{z\in\prod_{i\in\{1,\ldots,n\}} V_i}$ with finite support such that $x = \sum_{z\in\prod_{i\in\{1,\ldots,n\}} V_i} \alpha_z^{(x)} \cdot \delta_z$. Define now $\gamma: \mathcal{F}\left(\prod_{i\in\{1,\ldots,n\}} V_i, F\right) \rightarrow U$ by $x \rightarrow \sum_{z\in\Pi_{i\in\{1,\ldots,n\}} V_i} \alpha_z^{(x)} \cdot \varphi(z)$ then we have that
   a. $(\boldsymbol{\gamma} \circ \boldsymbol{\delta} = \boldsymbol{\varphi})$ If $x \in \prod_{i\in\{1,\ldots,n\}} V_i$ then $\delta(x) = \delta_x$ so that $(\gamma \circ \delta)(x) = \gamma(\delta(x)) = 1 \cdot f(x) = \varphi(x)$

b. **(γ is linear)** as $\alpha \cdot x + \beta \cdot y = \alpha \cdot \sum_{z \in \prod_{i \in \{1, \ldots, n\}} V_i} \alpha_z^{(x)} + \beta \cdot \sum_{z \in \prod_{i \in \{1, \ldots, n\}} V_i} \alpha_z^{(y)} = \sum_{z \in \prod_{i \in \{1, \ldots, n\}} V_i} \left(\alpha \cdot \alpha_z^{(x)} + \beta \cdot \alpha_z^{(y)}\right)$ so that $\gamma(\alpha \cdot x + \beta \cdot y) = \sum_{z \in \prod_{i \in \{1, \ldots, n\}} V_i} \left(\alpha \cdot \alpha_z^{(x)} + \beta \cdot \alpha_z^{(y)}\right) \cdot \varphi(z) = \alpha \cdot \sum_{z \in \prod_{i \in \{1, \ldots, n\}} V_i} \alpha_z^{(x)} \cdot \varphi(z) + \beta \cdot \sum_{z \in \prod_{i \in \{1, \ldots, n\}} V_i} \alpha_z^{(y)} \cdot \varphi(z) = \alpha \cdot \gamma(x) + \beta \cdot \gamma(y)$

c. **($\mathcal{S}\left(\mathcal{N}\left(\prod_{i \in \{1, \ldots, n\}} V_i, F\right)\right) \subseteq \ker(\gamma)$)** If $z \in \mathcal{N}\left(\prod_{i \in \{1, \ldots, n\}} V_i, F\right)$ then there $\exists i \in \{1, \ldots, n\}$, $\exists (x_1, \ldots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_n) \in \prod_{i \in \{1, \ldots, n\}} V_i$ and a $\alpha, \beta \in F$ such that $z = \delta_{(x_1, \ldots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_n)} - \alpha \cdot \delta_{(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)} - \beta \cdot \delta_{(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)}$ so that $\gamma(z) = 1 \cdot \varphi(x_1, \ldots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_n) - \alpha \cdot \varphi(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) - \beta \cdot \varphi(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) \underset{\varphi \text{ is multilinear}}{=} 0$. If now $z \in \mathcal{S}\left(\mathcal{N}\left(\prod_{i \in \{1, \ldots, n\}} V_i, F\right)\right)$ then there exists a $\{\alpha_x\}_{x \in \mathcal{N}(\prod_{i \in \{1, \ldots, n\}} V_i, F)}$ with finite support such that $z = \sum_{x \in \mathcal{N}(\prod_{i \in \{1, \ldots, n\}} V_i, F)} \alpha_x \cdot x$ so that we have $\gamma(x) = \gamma\left(\sum_{x \in \mathcal{N}(\prod_{i \in \{1, \ldots, n\}} V_i, F)} \alpha_x \cdot x\right) \underset{\gamma \text{ is linear}}{=} \sum_{x \in \mathcal{N}(\prod_{i \in \{1, \ldots, n\}} V_i, F)} \alpha_x \cdot \gamma(x) \underset{x \in \mathcal{N}(\prod_{i \in \{1, \ldots, n\}} V_i, F) \Rightarrow \gamma(x) = 0}{=} 0$ so we have $\mathcal{S}\left(\mathcal{N}\left(\prod_{i \in \{1, \ldots, n\}} V_i, F\right)\right) \subseteq \ker(\gamma)$.

Using 10.223 and (b,c) we have then the existence of a linear function h: $\mathcal{F}\left(\prod_{i \in \{1, \ldots, n\}} V_i, F\right) / \mathcal{S}\left(\mathcal{N}\left(\prod_{i \in \{1, \ldots, n\}} V_i, F\right)\right) = V_1 \otimes \ldots \otimes V_n \rightarrow U$ such that $\gamma = h \circ \pi$. Using (a) we have then $\varphi \underset{(a)}{=} \gamma \circ \delta = (h \circ \pi) \circ \delta = h \circ (\pi \circ \delta) \underset{\otimes = \pi \circ \delta}{=} h \circ \otimes$ □

We prove now some useful theorems about tensor products.

Theorem 10.383. *If $\{V_i\}_{i \in \{1, \ldots, n\}}$ is a finite non empty family of vector spaces over a field F with characterization 0, $\langle V_1 \otimes \ldots \otimes V_n, \otimes \rangle$ is a tensor product and $(x_1, \ldots, x_n) \in \prod_{i \in \{1, \ldots, n\}} V_i$ then we have $x_1 \otimes \ldots \otimes x_n = 0$ if and only if $\exists i \in \{1, \ldots, n\}$ such that $x_i = 0$*

Proof.

1. **($\Rightarrow$)** If $x_1 \otimes \ldots \otimes x_n = 0$ and assume that $\forall i \in \{1, \ldots, n\}$ we have $x_i = 0$ then by 10.176 there exists $\forall i \in \{1, \ldots, n\}$ a $L_i \in \mathrm{Hom}(V_i, F)$ such that $L_i(x_i) = 1$ then we have by 10.231 that $\prod_{i \in \{1, \ldots, n\}} L_i \in \mathrm{Hom}(V_1, \ldots, V_n, F)$ then there exists a linear mapping $h: V_1 \otimes \ldots \otimes V_n \rightarrow F$ such that $\left(\prod_{i \in \{1, \ldots, n\}} L_i\right) = h \circ \otimes$ and thus $0 = h(0) = h(x_1 \otimes \ldots \otimes x_n) = (h \circ \otimes)(x_1, \ldots, x_n) = \left(\prod_{i \in \{1, \ldots, n\}} L_i\right)(x_1, \ldots, x_n) = \prod_{i \in \{1, \ldots, n\}} L_i(x_i) = \prod_{i \in \{1, \ldots, n\}} 1 = 1$ giving that $1 = 0$ which is a contradiction as F is of characterization 0. So $\exists i \in \{1, \ldots, n\}$ such that $x_i = 0$

2. **($\Leftarrow$)** If $i \in \{1, \ldots, n\}$ is such that $x_i = 0$ then $x_i = 0 \cdot x_i$ so that we have $(x_1 \otimes, \ldots, \otimes x_n) = \otimes(x_1, \ldots, x_n) = \otimes(x_1, \ldots, x_{n-1}, 0 \cdot x_i, x_i, \ldots, x_n) = 0 \cdot (x_1, \ldots, x_n) = 0$ □

Next we consider the problem of defining a tensor product of sub spaces $W_i \subseteq V_i$ in such a way that $W_1 \otimes \ldots . \otimes W_n \subseteq V_1 \otimes \ldots V_n$, this is solved in the next theorem

Theorem 10.384. *Let $\{V_i\}_{i \in \{1, \ldots, n\}}$ a finite non empty family of vector spaces over a field F and let $\{W_i\}_{i \in \{1, \ldots, n\}}$ a family of sub spaces [$\forall i \in \{1, \ldots, .n\}$ we have W_i is a sub space of V_i]. If $\langle P, \nu \rangle$ is a tensor product of $\{V_i\}_{i \in \{1, \ldots, n\}}$ then if we define $\langle Q, \mu \rangle$ by (note that $\prod_{i \in \{1, \ldots, n\}} W_i \subseteq \prod_{i \in \{1, \ldots, n\}} V_i$)*

1. $Q = \mathcal{S}\left(\mu\left(\prod_{i \in \{1, \ldots, n\}} W_i\right)\right)$
2. $\mu = \nu_{|\prod_{i \in \{1, \ldots, n\}} W_I} : \prod_{i \in \{1, \ldots, n\}} W_i \rightarrow Q$

then we have that $\langle Q, \mu \rangle$ is a tensor product such that $Q \subseteq P$ and trivially by the definition of μ that $\forall (x_1, \ldots, x_n) \in \prod_{i \in \{1, \ldots, n\}} W_i$ we have $\mu(x_1, \ldots, x_n) = \nu(x_1, \ldots, x_n)$.

Proof. If $i \in \{1, \ldots, n\}$ then if $(x_1, \ldots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_n) \in \prod_{i \in \{1, \ldots, n\}} W_i$ then $\mu(x_1, \ldots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_n) = \nu(x_1, \ldots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_n) \underset{\nu \,\text{is multilinear}}{=} \alpha \cdot \nu(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) + \beta \cdot \nu(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) = \alpha \cdot \mu(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) + \beta \cdot \mu(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)$, so we have

$$\mu \in \operatorname{Hom}(W_1, \ldots, W_n, Q). \tag{10.142}$$

Further we have by definition that

$$\mathcal{S}\left(\mu\left(\prod_{i \in \{1, \ldots, n\}} W_i\right)\right) = Q. \tag{10.143}$$

Now if U is a vector space over the field F and $\psi \in \operatorname{Hom}(W_1, \ldots, W_n; U)$ then using 10.244 there exists a $\varphi \in \operatorname{Hom}(V_1, \ldots, V_n; U)$ such that $\varphi_{|\prod_{i \in \{1, \ldots, n\}} W_i} = \psi$. As $\langle P, \nu \rangle$ is a tensor product there exists a $h \in \operatorname{Hom}(P, U)$ such that $h \circ \nu = \varphi$, define now $k = h_{|Q}: Q \subseteq P \rightarrow U$ then we have $\forall (w_1, \ldots, w_n) \in \prod_{i \in \{1, \ldots, n\}} W_i$ that

$$\begin{array}{rcl}
(k \circ \mu)(w_1, \ldots, w_n) & = & k(\mu(w_1, \ldots, w_n)) \\
& \underset{\mu = \nu_{|\prod_{i \in \{1, \ldots, n\}} W_i}}{=} & k(\nu(w_1, \ldots, w_n)) \\
& \underset{\nu(\prod_{i \in \{1, \ldots, n\}} W_i) = Q \wedge k = h_{|Q}}{=} & h(\nu(w_1, \ldots, w_n)) \\
& \underset{h \circ \nu = \varphi}{=} & \varphi(w_1, \ldots, w_n) \\
& \underset{\varphi_{|\prod_{i \in \{1, \ldots, n\}} W_i} = \psi}{=} & \psi(w_1, \ldots, w_n)
\end{array}$$

proving that

$$k \circ \mu = \psi \tag{10.144}$$

Using 10.142, 10.143 and 10.144 we have that $\langle Q, \mu\rangle$ is indeed a tensor product. □

In other words the above theorem means using the $\otimes$ notation that if $\{V_i\}_{i\in\{1,\ldots,n\}}$ is a family of vector spaces with $\{W_i\}_{i\in\{1,\ldots,n\}}$ a family of sub spaces then if $\langle V_1\otimes\ldots\otimes V_n, \otimes\rangle$ is a tensor product of $\{V_i\}_{i\in\{1,\ldots,n\}}$ then the restriction of $\otimes$ defined by $\otimes_{|\prod_{i\in\{1,\ldots,n\}}W_i}: \prod_{i\in\{1,\ldots,n\}} W_i \rightarrow \otimes\left(\prod_{i\in\{1,\ldots,n\}} W_i\right) \underset{\text{definition}}{=} W_1\otimes\ldots\otimes W_n$ forms a tensor product noted by $\langle W_1\otimes\ldots\otimes W_n, \otimes\rangle$ such that $W_1\otimes\ldots\otimes W_n \subseteq V_1\otimes\ldots\otimes V_n$ and $\forall (x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}} W_i$ we have $x_1\otimes\ldots\otimes x_n\in W_1\otimes\ldots\otimes W_n\subseteq V_1\otimes\ldots\otimes V_n$

Theorem 10.385. *Let $\{V_i\}_{i\in\{1,\ldots,2\}}$ be vector spaces over a field F and let $z\in V_1\otimes V_2$. Then $\{k\in\mathbb{N}|\exists\{u_i\}_{i\in\{1,\ldots,k\}}\text{ a family in } V_1 \text{ and } \exists\{v_i\}_{i\in\{1,\ldots,k\}} \text{ such that } z=\sum_{i\in\{1,\ldots,k\}} u_i\otimes v_i\}\neq\emptyset$ and thus its minimum exist (see 4.52). If $k_z=\min\left(\{k\in\mathbb{N}|\exists\{u_i\}_{i\in\{1,\ldots,k\}}\text{ a family in } V_1 \text{ and } \exists\{v_i\}_{i\in\{1,\ldots,k\}} \text{ such that } z=\sum_{i\in\{1,\ldots,k\}} u_i\otimes v_i\}\right)$ then $\{u_i\}_{i\in\{1,\ldots,k\}}$, $\{v_i\}_{i\in\{1,\ldots,k\}}$ are linear independent sets.*

Proof. As by definition $\mathcal{S}(\otimes(V_1\times V_2))=V_1\otimes V_2$ we have that $\forall v\in V_1\otimes V_2$ there exists by 10.132 finite families $\{v_i'\otimes u_i\}_{i\in\{1,\ldots,n\}}$ in $\otimes(V_1\times V_2)$ and $\{\alpha_i\}_{i\in\{1,\ldots,n\}}$ in F such that $v=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot(v_i'\otimes u_i)=\sum_{i\in\{1,\ldots,n\}} v_i\otimes u_i$ where $v_i=\alpha_i\cdot v_i'\in V_1$ so that $\{k\in\mathbb{N}|\exists\{u_i\}_{i\in\{1,\ldots,k\}}\text{ a family in } V_1 \text{ and } \exists\{v_i\}_{i\in\{1,\ldots,k\}} \text{ such that } z=\sum_{i\in\{1,\ldots,k\}} u_i\otimes v_i\}\neq\emptyset$. Let now $k=\min\left(\{k\in\mathbb{N}|\exists\{u_i\}_{i\in\{1,\ldots,k\}}\text{ a family in } V_1 \text{ and } \exists\{v_i\}_{i\in\{1,\ldots,k\}} \text{ such that } z=\sum_{i\in\{1,\ldots,k\}} u_i\otimes v_i\}\right)$. If now $z=\sum_{i\in\{1,\ldots,k\}} u_i^{(1)}\otimes u_i^{(2)}$ then we have that:

1. If $\{u_i\}_{i\in\{1,\ldots,k\}}$ is dependent then there exists a $i_0\in\{1,\ldots,k\}$ such that $u_{i_0}=\sum_{i\in\{1,\ldots,k\}\setminus\{i_0\}} c_i\cdot u_i$ then

$$\begin{aligned}
z &= \sum_{i\in\{1,\ldots,k\}} u_i\otimes v_i\\
&= \sum_{i\in\{1,\ldots,k\}\setminus\{i_0\}} u_i\otimes v_i + u_{i_0}\otimes v_{i_0}\\
&= \sum_{i\in\{1,\ldots,k\}\setminus\{i_0\}} u_i\otimes v_i + \left(\sum_{i\in\{1,\ldots,k\}\setminus\{i_0\}} c_i\cdot u_i\right)\times v_{i_0}\\
&= \sum_{i\in\{1,\ldots,k\}\setminus\{i_0\}} u_i\otimes v_i + \sum_{i\in\{1,\ldots,k\}\setminus\{i_o\}} (c_i\cdot(u_i\otimes v_{i_0}))\\
&= \sum_{i\in\{1,\ldots,k\}\setminus\{i_0\}} (u_i\otimes v_i + c_i\cdot(u_i\otimes v_{i_0}))\\
&= \sum_{i\in\{1,\ldots,k\}\setminus\{i_0\}} (u_i\otimes(v_i+c_i\cdot v_{i_0}))
\end{aligned}$$

If we define now the bijection $b:\{1,\ldots,k-1\}\rightarrow\{1,\ldots,k\}\setminus\{i_0\}$ by $b(i)=\begin{cases} i \text{ if } i<i_0\\ i+1 \text{ if } i_0\leqslant i\leqslant k-1\end{cases}$ and $\{u_i'\otimes v_i'\}_{i\in\{1,\ldots,k-1\}}$ by $u_i'\otimes v_i'=u_{b_i}\otimes(v_{b_i}+c_{b_i}\times v_{i_0})$ then we have $v=\sum_{i\in\{1,\ldots,k-1\}} u_i'\otimes v_i'$ contradicting the minimality of k.

2. If $\{v_{i_i}\}_{i\in\{1,\ldots,k\}}$ is dependent then there exists a $i_0\in\{1,\ldots,k\}$ such that $v_{i_0}=\sum_{i\in\{1,\ldots,k\}\setminus\{i_0\}} c_i\cdot v_i$ then

$$\begin{aligned}
z &= \sum_{i\in\{1,\ldots,k\}} u_i\otimes v_i \\
&= \sum_{i\in\{1,\ldots,k\}\setminus\{i_0\}} u_i\otimes v_i + u_{i_0}\otimes v_{i_0} \\
&= \sum_{i\in\{1,\ldots,k\}\setminus\{i_0\}} u_i\otimes v_i + u_{i_0}\otimes\left(\sum_{i\in\{1,\ldots,k\}\setminus\{i_0\}} c_i\cdot v_i\right) \\
&= \sum_{i\in\{1,\ldots,k\}\setminus\{i_0\}} u_i\otimes v_i + \sum_{i\in\{1,\ldots,k\}\setminus\{i_o\}} (c_i\cdot(u_{i_0}\otimes v_i)) \\
&= \sum_{i\in\{1,\ldots,k\}\setminus\{i_0\}} (u_i\otimes v_i + c_i\cdot(u_{i_0}\otimes v_i)) \\
&= \sum_{i\in\{1,\ldots,k\}\setminus\{i_0\}} ((u_i + c_i\cdot u_{i_0})\otimes v_i)
\end{aligned}$$

If we define now the bijection $b\colon\{1,\ldots,k-1\}\rightarrow\{1,\ldots,k\}\setminus\{i_0\}$ by $b(i)=\begin{cases} i \text{ if } i<i_0 \\ i+1 \text{ if } i_0\leqslant i\leqslant k-1\end{cases}$ and $\{u_i'\otimes v_i'\}_{i\in\{1,\ldots,k-1\}}$ by $u_i'\otimes v_i'=(u_{b_i}+c_{b_i}\cdot u_{i_0})\otimes v_{b_i}$ then we have $v=\sum_{i\in\{1,\ldots,k-1\}} u_i'\otimes v_i'$ contradicting the minimality of k.

So we have proved that $\{u_i\}_{i\in\{1,\ldots,k\}}$ and $\{v_i\}_{i\in\{1,\ldots,k\}}$ are linear dependent. □

Theorem 10.386. *Let $\{V_i\}_{i\in\{1,\ldots,m+1\}}$ be vector spaces over a field F and $V_1\otimes\ldots\otimes V_{m+1}$ is a tensor product then for the tensor product $V_1\otimes\ldots\otimes V_m$ there exists a tensor product of $V_1\otimes\ldots\otimes V_m$ and V_{m+1} such that $V_1\otimes\ldots\otimes V_{m+1}=(V_1\otimes\ldots\otimes V_m)\otimes V_{m+1}$*

Proof. To avoid confusion in notation we write $\langle V_1\otimes\ldots\otimes V_{m+1},\nu_{m+1}\rangle$ for the tensor product of $\{V_i\}_{i\in\{1,\ldots,m+1\}}$ $\langle V_1\otimes\ldots\otimes V_m,\nu_m\rangle$ for the tensor product of $\{V_i\}_{i\in\{1,\ldots,m\}}$ Using 10.382 we have also a tensor product of $V_1\otimes\ldots\otimes V_m$ and V_{m+1} that we call $\langle(V_1\otimes\ldots\otimes V_m)\otimes V_{m+1},\nu\rangle$. To prove the theorem we must construct a new tensor product $\langle V_1\otimes\ldots\otimes V_{m+1},\tau\rangle$ of $V_1\otimes\ldots\otimes V_m$ and V_{m+1}. Given $v_{m+1}\in V_{m+1}$ define

$$\beta_{v_m}\in\operatorname{Hom}(V_1,\ldots,V_m,V_1\otimes\ldots\otimes V_{m+1}) \text{ by } \beta_{v_m+1}(v_1,\ldots,v_m)=\nu_{m+1}(v_1,\ldots,v_m,v_{m+1}) \tag{10.145}$$

which is indeed multilinear because of the multilinearity of the tensor product ν_{M+1}. Also because of the multilinearity of ν_{m+1} we have for $\alpha,\beta\in F$ and ν_{m+1}, $\nu_{m+1}'\in V_{m+1}$ that

$$\beta_{\alpha\cdot\nu_{m+1}+\beta\cdot\nu_{m+1}'}=\alpha\cdot\beta_{\nu_{m+1}}+\beta\cdot\beta_{\nu_{m+1}'} \tag{10.146}$$

By the definition of the tensor product $\langle V_1 \otimes \ldots \otimes V_m, \nu_m \rangle$ we have the existence of a linear mapping

$$g_{v_m} \in \operatorname{Hom}(V_1 \otimes \ldots \otimes V_m, V_1 \otimes \ldots \otimes V_{m+1})$$

$$g_{v_{m+1}} \circ \nu_m = \beta_{\nu_m} \Rightarrow g_{v_{m+1}}(\nu_m(v_1, \ldots, v_m)) = \nu_{m+1}(v_1, \ldots, v_m, v_{m+1}) \tag{10.147}$$

Take now the multilinear mapping φ

$$\varphi \in \operatorname{Hom}(V_1, \ldots, V_{m+1}, (V_1 \otimes \ldots \otimes V_m) \otimes V_{m+1}) \text{ defined by}$$

$$\varphi(v_1, \ldots, v_{m+1}) = \nu(\nu_m(v_1, \ldots, v_m), v_{m+1}) \tag{10.148}$$

Then by the definition of the tensor product $\langle V_1 \otimes \ldots \otimes V_{m+1}, \nu_{m+1} \rangle$ there exists a linear mapping

$$h \in \operatorname{Hom}(V_1 \otimes \ldots \otimes V_{m+1}, (V_1 \otimes \ldots \otimes V_m) \otimes V_{m+1})$$

$$h \circ \nu_{m+1} = \varphi \Rightarrow h(\nu_{m+1}(v_1, \ldots, v_{m+1})) = \nu(\nu_m(v_1, \ldots, v_m), v_{m+1}) \tag{10.149}$$

Define now the multilinear mapping ψ

$$\psi: (V_1 \otimes \ldots \otimes V_m) \times V_{m+1} \rightarrow V_1 \otimes \ldots \otimes V_{m+1} \text{ by } \psi(z, v_{m+1}) = g_{v_{m+1}}(z) \tag{10.150}$$

Then using 10.147 we have that ψ is multilinear so

$$\psi \in \operatorname{Hom}((V_1 \otimes \ldots \otimes V_m), V_{m+1}, V_1 \otimes \ldots \otimes V_{m+1}) \tag{10.151}$$

Using the definition of the tensor product $\langle (V_1 \otimes \ldots \otimes V_m) \otimes V_{m+1}, \nu \rangle$ there exist then a linear mapping

$$k \in \operatorname{Hom}((V_1 \otimes \ldots \otimes V_m) \otimes V_{m+1}, V_1 \otimes \ldots \otimes V_{m+1}) \text{ with } k \circ \nu = \psi$$

So we have if $z = \nu_m(v_1, \ldots, v_m)$ that

$$\begin{aligned} k(\nu(\nu_m(v_1, \ldots, v_m), v_{m+1})) &= k(\nu(z, v_{m+1})) \\ &= \psi(z, v_{m+1}) \\ &= g_{v_{m+1}}(z) \\ &= g_{v_{m+1}}(\nu_m(v_1, \ldots, v_m)) \\ &\underset{10.147}{=} \nu_{m+1}(v_1, \ldots, v_m, v_{m+1}) \end{aligned} \tag{10.152}$$

Consider then the linear mapping $k \circ h: V_1 \otimes \ldots \otimes V_{m+1} \rightarrow V_1 \otimes \ldots \otimes V_{m+1}$ then we have if $z \in V_1 \otimes \ldots \otimes V_{m+1}$ that $z = \sum_{i \in \{1, \ldots, r\}} \nu_{m+1}(v_{i,1}, \ldots, v_{i,m+1})$ and

$$\begin{aligned} (k \circ h)(z) &= k\left(h\left(\sum_{i \in \{1, \ldots, r\}} \nu_{m+1}(v_{i,1}, \ldots, v_{i,m+1})\right)\right) \\ &= \sum_{i \in \{1, \ldots, r\}} k(h(\nu_{m+1}(v_{i,1}, \ldots, v_{i,m+1}))) \\ &\underset{10.149}{=} \sum_{i \in \{1, \ldots, r\}} k(\nu(\nu_m(v_{i,1}, \ldots, v_{i,m}), v_{i,m+1})) \\ &\underset{10.152}{=} \sum_{i \in \{1, \ldots, r\}} \nu_{m+1}(v_{i,1}, \ldots, v_{i,m+1}) \\ &= z \end{aligned}$$

Proving that

$$k\circ h=1_{V_1\otimes\ldots\otimes V_{m+1}} \tag{10.153}$$

Likewise we have for the linear mapping $h\circ k$: $(V_1\otimes\ldots\otimes V_m)\otimes V_{m+1}\rightarrow(V_1\otimes\ldots\otimes V_m)\otimes V_{m+1}$ that if $z\in(V_1\otimes\ldots\otimes V_m)\otimes V_{m+1}$ then $z=\sum_{i\in\{1,\ldots,r\}}\nu\left(\sum_{j\in\{1,\ldots,l\}}\nu_m(v_{i,j,1},\ldots,v_{i,j,m}),v_{i,m+1}\right)$

$$\begin{aligned}
(h\circ k)(z) &= h\left(k\left(\sum_{i\in\{1,\ldots,r\}}\nu\left(\sum_{j\in\{1,\ldots,l\}}\nu_m(v_{i,j,1},\ldots,v_{i,j,m}),v_{i,m+1}\right)\right)\right)\\
&= \sum_{i\in\{1,\ldots,r\}}h\left(k\left(\nu\left(\sum_{j\in\{1,\ldots,l\}}\nu_m(v_{i,j,1},\ldots,v_{i,j,m}),v_{i,m+1}\right)\right)\right)\\
&\underset{\nu\,\text{is mulitlinear}}{=} \sum_{i\in\{1,\ldots,r\}}h\left(k\left(\sum_{j\in\{1,\ldots,l\}}\nu(\nu_m(v_{i,j,1},\ldots,v_{i,j,m}),v_{i,m+1})\right)\right)\\
&= \sum_{i\in\{1,\ldots,r\}}\left(\sum_{j\in\{1,\ldots,l\}}h(k(\nu(\nu_m(v_{i,j,1},\ldots,v_{i,j,m}),v_{i,m+1})))\right)\\
&\underset{10.152}{=} \sum_{i\in\{1,\ldots,r\}}\left(\sum_{j\in\{1,\ldots,l\}}h(\nu_{m+1}(\nu_{m+1}(v_{i,j,1},\ \ldots,\ v_{i,j,m},\right.\\
&\qquad \left. v_{i,m+1})))\right)\\
&\underset{10.149}{=} \sum_{i\in\{1,\ldots,r\}}\nu\left(\sum_{j\in\{1,\ldots,l\}}\nu(\nu_m(v_{i,j,1},\ldots,v_{i,j,m}),v_{i,m+1})\right)\\
&= z
\end{aligned}$$

Proving that

$$h\circ k=1_{(V_1\otimes\ldots\otimes V_m)\otimes V_{m+1}} \tag{10.154}$$

Define now

$$\tau\colon(V_1\otimes\ldots\otimes V_m)\times V_{m+1}\rightarrow V_1\otimes\ldots\otimes V_{m+1}\,\text{by}\,\tau(z,v_{m+1})=k(\nu(z,v_{m+1})) \tag{10.155}$$

Then by the linearity of k and the multilinearity of ν we have that τ is multilinear or

$$\tau\in\operatorname{Hom}((V_1\otimes\ldots\otimes V_m),V_{m+1},V_1\otimes\ldots\otimes V_{m+1}) \tag{10.156}$$

We finish by proving that $\langle V_1\otimes\ldots\otimes V_{m+1},\tau\rangle$ is a tensor product of $V_1\otimes\ldots\otimes V_m$ and V_{m+1}.

1. $V_1\otimes\ldots\otimes V_{m+1}$ is a vector space space over the field F (by the definition of $\langle V_1\otimes\ldots\otimes V_{m+1},\nu_{m+!}\rangle$
2. By 10.168 we have $\tau\in\operatorname{Hom}((V_1\otimes\ldots\otimes V_m),V_{m+1},V_1\otimes\ldots\otimes V_{m+1})$
3. We have as $\tau(V_1\otimes\ldots\otimes V_m\times V_{m+1})\subseteq V_1\otimes\ldots\otimes V_{m+1}$ that $\mathcal{S}(\tau(V_1\otimes\ldots\otimes V_m\times V_m))\subseteq\mathcal{S}(V_1\otimes\ldots\otimes V_{m+1})=V_1\otimes\ldots\otimes V_{m+1}\Rightarrow\tau(V_1\otimes\ldots\otimes V_m\times V_{m+1})\subseteq V_1\otimes\ldots\otimes V_{m+1}$. If $v\in V_1\otimes\ldots\otimes V_{m+1}$ then $v=\sum_{i\in\{1,\ldots,l\}}\nu_{m+1}(v_1,\ \ldots,\ v_{m+1})\underset{10.153}{=}\sum_{i\in\{1,\ldots,l\}}k(h(\nu_{m+1}(v_1,\ \ldots,\ v_{m+1})))\underset{10.149}{=}\sum_{i\in\{1,\ldots,l\}}k(\nu(\nu_m(v_1,\ \ldots,\ v_m),$

$v_{m+1}))\underset{10.167}{=}\sum_{i\in\{1,\ldots,l\}}\tau(\nu_m(v_1,\ldots,v_m),v_{m+1})\in\mathcal{S}(\tau(V_1\otimes\ldots\otimes V_m\times V_{m+1}))$ proving that $V_1\otimes\ldots\otimes V_{m+1}\subseteq\mathcal{S}(\tau(V_1\otimes\ldots\otimes V_m\times V_{m+1}))$. This proves that $\mathcal{S}(\tau(V_1\otimes\ldots\otimes V_m\times V_{m+1}))=V_1\otimes\ldots\otimes V_{m+1}$

4. If $\theta\in\mathrm{Hom}(V_1\otimes\ldots\otimes V_m,V_{m+1},U)$ then by the definition of $\langle(V_1\otimes\ldots\otimes V_m)\otimes V_{m+1},\nu\rangle$ there exists a $\xi\in\mathrm{Hom}((V_1\otimes\ldots\otimes V_m)\otimes V_{m+1},U)$ such that $\xi\circ\nu=\theta$. Define now $\lambda=\xi\circ h: V_1\otimes\ldots\otimes V_{m+1}\rightarrow U$ which is linear because ξ and h are linear, so that $\lambda\in\mathrm{Hom}(V_1\otimes\ldots\otimes V_{m+1},U)$. Further we have if $(z,v_{m+1})\in V_1\otimes\ldots\otimes V_m\times V_{m+1}$

$$\begin{aligned}(\lambda\circ\tau)(z,v_{m+1}) &= \lambda(\tau(z,v_{m+1}))\\ &= \lambda(k(\nu(z,v_{m+1})))\\ &\underset{10.154}{=} \xi(h(k(\nu(z,v_{m+1}))))\\ &= \xi(\nu(z,v_{m+1}))\\ &= \theta(z,v_{m+1})\end{aligned}$$

proving that $\lambda\circ\tau=\theta$ □

Lemma 10.387. *Let $\{X_i\}_{i\in\{1,\ldots,n\}}$ be a finite family of finite sets then $\prod_{i\in\{1,\ldots,n\}}X_i$ is a finite set and $\#\left(\prod_{i\in\{1,\ldots,n\}}X_i\right)=\prod_{i\in\{1,\ldots,n\}}\#(X_i)$*

Proof. First note that if $m>2$ then we have that $\prod_{i\in\{1,\ldots,m\}}X_i$ is bijective with $\left(\prod_{i\in\{1,\ldots,m-1\}}X_i\right)\times X_m$ with $b:\prod_{i\in\{1,\ldots,m\}}X_i\rightarrow\left(\prod_{i\in\{1,\ldots,m-1\}}X_i\right)\times X_m$ where $b(x_1,\ldots,x_m)=((x_1,\ldots,x_{m-1}),x_m)$

1. **(injectivity)** If $b(x_1,\ldots,x_m)=b(y_1,\ldots,y_m)\Rightarrow((x_1,\ldots,x_{m-1}),x_m)=((y_1,\ldots,y_{m-1}),y_m)\Rightarrow x_m=y_m\wedge(x_1,\ldots,x_{m-1})=(y_1,\ldots,y_{m-1})\Rightarrow(x_1,\ldots,x_m)=(y_1,\ldots,y_m)$
2. **(surjectivity)** If $((x_1,\ldots,x_{m-1}),y)\in\left(\prod_{i\in\{1,\ldots,m-1\}}X_i\right)\times X_m$ then $b(x_1,\ldots,x_{m-1},y)=((x_1,\ldots,x_{m-1}),y)$

Next we proceed by induction so let $S=\left\{n\in\mathbb{N}|\text{if }\{X_i\}_{i\in\{1,\ldots,n\}}\text{ be a finite family of finite sets then }\prod_{i\in\{1,\ldots,n\}}\text{ is finite with }\#\left(\prod_{i\in\{1,\ldots,n\}}X_i\right)=\prod_{i\in\{1,\ldots,n\}}\#(X_i)\right\}$ then we have

1. **($n=1$)** then $\prod_{i\in\{1,\ldots,1\}}X_i$ is bijective with X_1 by using the bijection $b:\prod_{i\in\{1,\ldots,1\}}X_i\rightarrow X_1$ where $b(x)=x(1)$
   a. **(injectivity)** If $b(x)=b(y)=x(1)=y(1)\Rightarrow\forall i\in\{1,\ldots,1\}$ we have $x(i)=y(i)\Rightarrow x=y$
   b. **(surjectivity)** If $x\in X_1$ take then $x'\in\prod_{i\in\{1,\ldots,1\}}X_i$ by $\forall i\in\{1,\ldots,1\}$ we have $x'(i)=x'(1)=x$ so that $b(x')=x'(1)=x$

   So we have proved that $\#\left(\prod_{i\in\{1,\ldots,1\}}X_i\right)=\#(X_1)=\prod_{i\in\{1,\ldots,1\}}\#(X_i)$ proving that $1\in S$
2. **($n\in S$)** then for $n+1$ we have that $\prod_{i\in\{1,\ldots,n+1\}}X_i$ is bijective with $\left(\prod_{i\in\{1,\ldots,n\}}X_i\right)\times X_{n+1}$ which by 5.44 is finite and for which we have $\#\left(\left(\prod_{i\in\{1,\ldots,n\}}X_i\right)\times X_{n+1}\right)=\#\left(\prod_{i\in\{1,\ldots,n\}}X_i\right)\cdot\#(X_{n+1})$. So we have $\#\left(\prod_{i\in\{1,\ldots,n+1\}}X_i\right)=\#\left(\prod_{i\in\{1,\ldots,n\}}X_i\right)\cdot$

$\#(X_{n+1}) \underset{n\in S}{=} \left(\prod_{i\in\{1,\ldots,n\}} \#(X_i)\right) \cdot \#(X_{n+1}) = \prod_{i\in\{1,\ldots,n+1\}} \#(X_i)$ and thus $n+1\in S$

Using mathematical induction we have then proved our theorem. □

Definition 10.388. *Let* $\{n_i\}_{i\in\{1,\ldots,k\}}$ *be a family in* $\mathbb{N}$ *then* $\Gamma(n_1,\ldots,n_k) = \prod_{i\in\{1,\ldots,k\}} \{1,\ldots,n_i\}$. *Using the above lemma we have then that* $\#(\Gamma(k_1,\ldots,k_m)) = \prod_{i\in\{1,\ldots,m\}} k_i$

Example 10.389. $\Gamma(2,3) = \{(1,1),(1,2),(1,3),(2,1),(2,2),(2,3)\}$

Theorem 10.390. *Let* $\{V_i\}_{i\in\{1,\ldots,m\}}$ *be vector spaces over a field* F, V *a vector space over the field* F *and* $\varphi: \prod_{i\in\{1,\ldots,m\}} V_i \to V$ *a multilinear mapping,* $\{k_i\}_{i\in\{1,\ldots,m\}}$ *a family in* $\mathbb{N}$ *and* $\forall i\in\{1,\ldots,m\}$ *a family* $\{w_{i,j}\}_{j\in\{1,\ldots,k_i\}}$ *in* V_i *then* $\varphi\left(\sum_{j\in\{1,\ldots,k_1\}} w_{1,j},\ldots,\sum_{j\in\{1,\ldots,k_m\}} w_{m,j}\right) = \sum_{\gamma\in\Gamma} \varphi(w_{1,\gamma(1)},\ldots,w_{m,\gamma(m)})$

Proof. We prove this by induction on let $S = \{m \in \mathbb{N} \mid \text{the assertion of the theorem is true}\}$ then we have

1. $(\boldsymbol{m=1})$ In this case we have that $\Gamma(k_1) = \{1,\ldots,k_1\}^{\{1\}} = \{f: \{1\} \to \{1,\ldots,k_1\}\}$ if $\gamma \in \Gamma(k_1)$ and we can trivially make a bijection from $\tau: \{1,\ldots,k_1\} \to \Gamma(k_1)$ by $i \to \tau(i)$ where $(\tau(i))(1) = i$

   a. **(injectivity)** If $\tau(i) = \tau(j) \Rightarrow (\tau(i))(1) = (\tau(j))(1) \Rightarrow i = j$

   b. **(surjectivity)** If $\gamma \in \Gamma(k_1)$ take then $i = \gamma(1) \in \{1,\ldots,k_1\}$ then we have that $\forall j \in \{1\}$ that $(\tau(i))(j) = (\tau(i))(1) = i = \gamma(1) = \gamma(j) \Rightarrow \gamma = \tau(i)$

   As $\varphi\left(\sum_{j\in\{1,\ldots,k_1\}} w_{1,j}\right) \underset{\varphi\,\text{is multilinear}}{=} \sum_{j\in\{1,\ldots,k_1\}} \varphi(w_{1,j}) = \sum_{j\in\{1,\ldots,k_1\}} \varphi(w_{1,\tau(j)(1)}) \underset{10.44}{=} \sum_{\gamma\in\Gamma(k_1)} \varphi(w_{1,\gamma(1)})$ proving that $1 \in S$

2. $(\boldsymbol{m\in S})$ then if we take $m+1$ and $\varphi: \prod_{i\in\{1,\ldots,m+1\}} V_i \to V$ then if we define (see 10.238) $\varphi_{\ldots\sum_{j\in\{1,\ldots,k_{m+1}\}} w_{m+1,j}}: \prod_{i\in\{1,\ldots,m\}} V_i \to V$ by $\varphi_{\ldots\sum_{j\in\{1,\ldots,k_{m+1}\}} w_{m+1,j}}(x_1,\ldots,x_m) = \varphi\left(x_1,\ldots,x_m,\sum_{j\in\{1,\ldots,k_{m+1}\}} w_{m+1,j}\right)$ then we have that by 10.238 that $\varphi_{\ldots\sum_{j\in\{1,\ldots,k_{m+1}\}} w_{m+1,j}}$ is multilinear and as $m\in S$ we have then that $\varphi\left(\sum_{j\in\{1,\ldots,k_1\}} w_{1,j},\ldots,\sum_{j\in\{1,\ldots,k_{m+1}\}} w_{m+1,j}\right) = \varphi_{\ldots\sum_{j\in\{1,\ldots,k_{m+1}\}} w_{m+1,j}}\left(\sum_{j\in\{1,\ldots,k_1\}} w_{1,j},\ldots,\sum_{j\in\{1,\ldots,k_m\}} w_{m,j}\right) = \sum_{\gamma\in\Gamma(k_1,\ldots,k_m)} \varphi_{\ldots\sum_{j\in\{1,\ldots,k_{m+1}\}} w_{m+1,j}}(w_{1,\gamma(1)},\ldots,w_{m,\gamma(m)}) = \sum_{\gamma\in\Gamma(k_1,\ldots,k_m)} \varphi\left(w_{1,\gamma(1)},\ldots,w_{m,\gamma(m)},\sum_{j\in\{1,\ldots,k_{m+1}\}} w_{m+1,j}\right) \underset{\varphi\,\text{is multilinear}}{=} \sum_{\gamma\in\Gamma(k_1,\ldots,k_m)} \left(\sum_{j\in\{1,\ldots,k_{m+1}\}} \varphi(w_{1,\gamma(1)},\ldots,w_{m,\gamma(m)},w_{m+1,j})\right)$ or

$$\varphi\left(\sum_{j\in\{1,\ldots,k_1\}} w_{1,j},\ldots,\sum_{j\in\{1,\ldots,k_{m+1}\}} w_{m+1,j}\right) = \sum_{\gamma\in\Gamma(k_1,\ldots,k_m)} \left(\sum_{j\in\{1,\ldots,k_{m+1}\}} \varphi(w_{1,\gamma(1)},\ldots,w_{m,\gamma(m)},w_{m+1,j})\right) \quad (10.157)$$

Define now $\chi: \Gamma(k_1, \ldots, k_m) \times \{1, \ldots, k_{m+1}\} \rightarrow \Gamma(k_1, \ldots, k_{m+1})$ by $(\gamma, i) \rightarrow \chi(\gamma, i)$ by $\forall j \in \{1, \ldots, m+1\}$ we have $\chi(\gamma, i)(j) = \left\{ \begin{smallmatrix} \gamma(j) \text{ if } j \in \{1, \ldots, m\} \\ i \text{ if } j = m+1 \end{smallmatrix} \right.$. We prove now that χ is a bijection

1. **(injectivity)** If $\chi(\gamma, i) = \chi(\gamma', i')$ then $\forall j \in \{1, \ldots, m\}$ we have $\chi(\gamma, i)(j) = \chi(\gamma', i')(j) \Rightarrow \gamma(j) = \gamma'(j)$ and thus we have $\gamma = \gamma'$, also we have $\chi(\gamma, i)(m+1) = \chi(\gamma', i')(m+1) \Rightarrow i = i'$. So finally we must have $(\gamma, i) = (\gamma', i')$ proving injectivity.

2. **(surjectivity)** If $\gamma \in \Gamma(k_1, \ldots, k_{m+1})$ taken then $\gamma_{|\{1, \ldots m\}} \in \Gamma(k_1, \ldots, k_m)$ and $\gamma(m+1) \in \{1, \ldots, k_{m+1}\}$ then we have if $j \in \{1, \ldots, m\}$ that $\chi(\gamma_{|\{1, \ldots, m\}}, \gamma(m+1))(j) = \gamma_{|\{1, \ldots, m\}}(j) = \gamma(j)$ and if $j = m+1$ then we have $\chi(\gamma_{|\{1, \ldots, m\}}, \gamma(m+1)) = \gamma(m+1)$. This proves that $\chi(\gamma_{|\{1, \ldots, m\}}, \gamma(m+1)) = \gamma$ and thus surjectivity. □

Using χ we have then that $\sum_{(\gamma, i) \in \Gamma(k_1, \ldots, k_m) \times \{1, \ldots, k_{m+1}\}} \varphi(w_{1,\gamma(1)}, \ldots, w_{m,\gamma(m)}, w_{m+1,i}) = \sum_{(\gamma, i) \in \Gamma(k_1, \ldots, k_m) \times \{1, \ldots, k_{m+1}\}} \varphi(w_{1,\chi(\gamma,i)(1)}, \ldots, w_{m,\chi(\gamma,i)(m)}, w_{m+1,\chi(\gamma,i)(m+1)}) \underset{10.44}{=} \sum_{\gamma \in \Gamma(k_1, \ldots, k_{m+1})} \varphi(w_{1,\gamma(1)}, \ldots, w_{m+1,\gamma(m+1)})$ or

$$\sum_{(\gamma, i) \in \Gamma(k_1, \ldots, k_m) \times \{1, \ldots, k_{m+1}\}} \varphi(w_{1,\gamma(1)}, \ldots, w_{m,\gamma(m)}, w_{m+1,i}) = \sum_{\gamma \in \Gamma(k_1, \ldots, k_{m+1})} \varphi(w_{1,\gamma(1)}, \ldots, w_{m+1,\gamma(m+1)}) \tag{10.158}$$

Given now a $\gamma \in \Gamma(k_1, \ldots, k_m)$ construct then the trivially bijective function ρ_γ: $\{\gamma\} \times \{1, \ldots, k_{m+1}\} \rightarrow \{1, \ldots, k_{m+1}\}$ by $(\gamma, i) \rightarrow \rho_\gamma(\gamma, i) = i$

Now as $\Gamma(k_1, \ldots, k_m) \times \{1, \ldots, k_{m+1}\} = \bigcup_{\gamma \in \Gamma(k_1, \ldots, k_m)} \{\gamma\} \times \{1, \ldots, k_{m+1}\}$ and also that trivially the union is pairwise disjoint then we have by 10.45 and that $\sum_{(\gamma, i) \in \Gamma(k_1, \ldots, k_m) \times \{1, \ldots, k_{m+1}\}} \varphi(w_{1,\gamma(1)}, \ldots, w_{m,\gamma(m)}, w_{m+1,i}) = \sum_{\gamma \in \Gamma(k_1, \ldots, k_m)} \left(\sum_{(\kappa, i) \in \{\gamma\} \times \{1, \ldots, k_{m+1}\}} \varphi(w_{1,\kappa(1)}, \ldots, w_{m,\kappa(m)}, w_{m+1,i})\right) = \sum_{\gamma \in \Gamma(k_1, \ldots, k_m)} \left(\sum_{(\kappa, i) \in \{\gamma\} \times \{1, \ldots, k_{m+1}\}} \varphi(w_{1,\gamma(1)}, \ldots, w_{m,\gamma(m)}, w_{m+1,i})\right) = \sum_{\gamma \in \Gamma(k_1, \ldots, k_m)} \left(\sum_{(\kappa, i) \in \{\gamma\} \times \{1, \ldots, k_{m+1}\}} \varphi(w_{1,\gamma(1)}, \ldots, w_{m,\gamma(m)}, w_{m+1,\rho_\gamma(\kappa,i)})\right) \underset{10.44}{=} \sum_{\gamma \in \Gamma(k_1, \ldots, k_m)} \left(\sum_{j \in \{1, \ldots, k_{m+1}\}} \varphi(w_{1,\gamma(1)}, \ldots, w_{m,\gamma(m)}, w_{m+1,j})\right)$ or

$$\sum_{(\gamma, i) \in \Gamma(k_1, \ldots, k_m) \times \{1, \ldots, k_{m+1}\}} \varphi(w_{1,\gamma(1)}, \ldots, w_{m,\gamma(m)}, w_{m+1,i}) = \sum_{\gamma \in \Gamma(k_1, \ldots, k_m)} \left(\sum_{j \in \{1, \ldots, k_{m+1}\}} \varphi(w_{1,\gamma(1)}, \ldots, w_{m,\gamma(m)}, w_{m+1,j})\right) \tag{10.159}$$

Finally we have by 10.157,10.158 and the above 10.159 that $\varphi\left(\sum_{j \in \{1, \ldots, k_1\}} w_{1,j}, \ldots, \sum_{j \in \{1, \ldots, k_{m+1}\}} w_{m+1,j}\right) = \sum_{\gamma \in \Gamma(k_1, \ldots, k_{m+1})} \varphi(w_{1,\gamma(1)}, \ldots, w_{m+1,\gamma(m+1)})$ proving that $m+1 \in S$.

By mathematical induction we have then proved the theorem.

Lemma 10.391. *Let $\{V_i\}_{i\in\{1,\ldots,n\}}$ be vector spaces over a field F such that $\forall i\in\{1,\ldots,n\}$ there exists a family $\{W_{i,j}\}_{j\in\{1,\ldots,k_i\}}$ of sub spaces of V_i such that $V_i=\sum_{j\in\{1,\ldots,k_i\}} W_{i,j}$ (see 10.357) then we have that $V_1\otimes\ldots\otimes V_n=\sum_{\gamma\in\Gamma(k_1,\ldots,k_n)}(W_{1,\gamma(1)}\otimes\ldots\otimes W_{m,\gamma(m)})$ (where the $\otimes$ on the $W_{i,j}$ is defined by 10.384 based on $\otimes$ of the $V_i's$)*

Proof. First as $W_{1,\gamma(1)}\otimes\ldots\otimes W_{n,\gamma(n)}\subseteq V_1\otimes\ldots\otimes V_n$ (see 10.384) $\forall\gamma\in\Gamma(k_1,\ldots,k_n)$ we have if $v\in\sum_{\gamma\in\Gamma(k_1,\ldots,k_n)}(W_{1,\gamma(1)}\otimes\ldots\otimes W_{n,\gamma(n)})$ then $v=\sum_{\gamma\in\Gamma(k_1,\ldots,k_n)} v_\gamma$ where $\forall\gamma\in\Gamma(k_1,\ldots,k_n)$ we have $v_\gamma\in W_{1,\gamma(1)}\otimes\ldots\otimes W_{n,\gamma(n)}\subseteq V_1\otimes\ldots\otimes V_n$ so that $v=\sum_{\gamma\in\Gamma(k_1,\ldots,k_n)} v_\gamma\in V_1\otimes\ldots\otimes V_n$ proving that

$$\sum_{\gamma\in\Gamma(k_1,\ldots,k_n)}(W_{1,\gamma(1)}\otimes\ldots\otimes W_{n,\gamma(m)})\subseteq V_1\otimes\ldots\otimes V_m. \tag{10.160}$$

Now if $(v_1,\ldots,v_n)\in\prod_{i\in\{1,\ldots,n\}} V_i$ then $\forall i\in\{1,\ldots,n\}$ there exists a $\{w_{i,j}\}_{j\in\{1,\ldots,k_i\}}$ such that $v_i=\sum_{j\in\{1,\ldots,k_i\}} w_{i,j}$. As $\otimes$ is multilinear we have by 10.390 that $v_1\otimes\ldots\otimes v_m=\otimes(v_1,\ldots,v_n)=\sum_{\gamma\in\Gamma(k_1,\ldots,k_n)}\otimes(v_{1,\gamma(1)},\ldots,v_{n,\gamma(n)})=\sum_{\gamma\in\Gamma(k_1,\ldots,k_m)} w_{1,\gamma(1)}\otimes\ldots\otimes w_{m,\gamma(m)}\in\sum_{\gamma\in\Gamma(k_1,\ldots,k_n)}(W_{1,\gamma(1)}\otimes\ldots\otimes W_{m,\gamma(m)})$ proving that $\otimes\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)\subseteq\sum_{\gamma\in\Gamma(k_1,\ldots,k_n)}(W_{1,\gamma(1)}\otimes\ldots\otimes W_{n.\gamma(n)})$. From $V_1\otimes\ldots\otimes V_n=\mathcal{S}\left(\otimes\left(\prod_{i\in\{1,\ldots,n\}} V_n\right)\right)\subseteq_{10.136}\mathcal{S}\left(\sum_{\gamma\in\Gamma(k_1,\ldots,k_n)}(W_{1,\gamma(1)}\otimes\ldots\otimes W_{n,\gamma(n)})\right)\underset{10.363\text{ and }10.134}{=}\sum_{\gamma\in\Gamma}(W_{1,\gamma(1)}\otimes\ldots\otimes W_{n,\gamma(n)})$ which together with 10.160 gives that

$$V_1\otimes\ldots\otimes V_n=\sum_{\gamma\in\Gamma(k_1,\ldots,k_n)}(W_{1,\gamma(1)}\otimes\ldots\otimes W_{n,\gamma(n)})$$ □

Lemma 10.392. *Let $\{V_i\}_{i\in\{1,2\}}$ be vector spaces over a field F (with characterization 0) and $V_1=W_1\oplus W_2$ (or $V_2=W_1\oplus W_2$) then $V_1\otimes V_2=(W_1\otimes V_2)\oplus(W_2\otimes V_2)$ (or $V_1\otimes V_2=(V_1\otimes W_1)\oplus(V_2\otimes W_2)$)*

Proof.

1. **($V_1=W_1\oplus W_2$)** By 10.391 we have already proved that $V_1\otimes V_2=(W_1\otimes V_2)+(W_2\otimes V_2)$, so the only thing left to prove is that $(W_1\otimes V_2)\bigcap(W_2\otimes V_2)=\{0\}$. We do this by contradiction, so assume that $x\in(W_1\otimes V_2)\bigcap(W_2\otimes V_2)$ with $x\neq 0$ then using 10.378 there exists families $\{u_i\}_{i\in\{1,\ldots,k\}}$ of elements of W_1, $\{v_i\}_{i\in\{1,\ldots,r\}}$ of elements of W_2 and $\{x_i\}_{i\in\{1,\ldots,k\}}$, $\{y_i\}_{i\in\{1,\ldots,r\}}$ of elements of V_2 such that

$$\begin{aligned} x &= \sum_{i\in\{1,\ldots,k\}} u_i\otimes x_i\in W_1\otimes V_2\\ &= \sum_{i\in\{1,\ldots,r\}} v_i\otimes y_i\in W_2\otimes V_2 \end{aligned} \tag{10.161}$$

in which r, k are minimal. Using 10.385 we have then that $\{u\}_{i \in \{1,\ldots,k\}}$, $\{v_i\}_{i \in \{1,\ldots,r\}}$, $\{x_i\}_{i \in \{1,\ldots,k\}}$, $\{y_i\}_{i \in \{1,\ldots,r\}}$ are linear independent sets. As $\{x_i\}_{i \in \{1,\ldots,k\}}$ is linear independent we have using 10.178 the existence of a linear mapping $h: V_2 \rightarrow F$ such that $h(x_1) = 1$ and $\forall i \in \{2, \ldots, k\}$ we have $h(x_i) = 0$. Choose a arbitrary linear mapping $f: V_1 \rightarrow F$ then we have that $f \cdot h: V_1 \times V_2 \rightarrow F$ is a multilinear mapping and thus by the definition of the tensor product there exists a unique linear $g: V_1 \otimes V_2 \rightarrow F$ such that

$$g \circ \otimes = f \cdot h \tag{10.162}$$

Using 10.161 we have then

$$\begin{aligned} g(x) &= g\left(\sum_{i \in \{1,\ldots,k\}} u_i \otimes x_i\right) \\ &= \sum_{i \in \{1,\ldots,k\}} g(u_i \otimes x_i) \\ &= \sum_{i \in \{1,\ldots,k\}} (g \circ \otimes)(u_i, x_i) \\ &= \sum_{i \in \{1,\ldots,k\}} f(u_i) \cdot h(x_i) \\ &= f(u_1) \end{aligned} \tag{10.163}$$

But we have also

$$\begin{aligned} g(x) &= g\left(\sum_{i \in \{1,\ldots,r\}} v_i \otimes y_i\right) \\ &= \sum_{i \in \{1,\ldots,r\}} g(v_i \otimes y_i) \\ &= \sum_{i \in \{1,\ldots,r\}} f(v_i) \cdot h(y_i) \\ &\underset{f is \text{linear}}{=} f\left(\sum_{i \in \{1,\ldots,r\}} h(y_i) \cdot v_i\right) \end{aligned}$$

so that we have for a arbitrary linear function $f: V_1 \rightarrow F$ that $f(u_1) = f\left(\sum_{i \in \{1,\ldots,r\}} h(y_i) \cdot v_i\right)$ which gives by 10.177 that $W_1 \ni u_1 = \sum_{i \in \{1,\ldots,r\}} h(y_i) \cdot v_i \in W_2$. This gives that $u_1 \in W_1 \bigcap W_2 = \{0\}$ or $u_1 = 0$ contradicting the linear Independence of $\{u_i\}_{i \in \{1,\ldots,k\}}$. So we conclude that our assumption of $x \neq 0$ is wrong and thus we must have that $(W_1 \otimes V_2) \bigcap (W_2 \otimes V_2) = \{0\}$.

2. $(\boldsymbol{V_2 = W_1 \oplus W_2})$ This is proved in a similar way as (1). By 10.391 we have already proved that $V_1 \otimes V_2 = (V_1 \otimes W_2) + (V_1 \otimes W_2)$, so the only thing left to prove is that $(V_1 \otimes W_1) \bigcap (V_1 \otimes W_2) = \{0\}$. We do this by contradiction, so assume that $x \in (V_1 \otimes W_1) \bigcap (V_1 \otimes W_2)$ with $x \neq 0$ then using 10.378 there

exists families $\{u_i\}_{i\in\{1,\ldots,k\}}$, $\{v_i\}_{i\in\{1,\ldots,r\}}$ of elements of V_1, $\{x_i\}_{i\in\{1,\ldots,k\}}$ a family in W_1 and $\{y_i\}_{i\in\{1,\ldots,r\}}$ a family in W_2 such that

$$\begin{aligned} x &= \sum_{i\in\{1,\ldots,k\}} u_i\otimes x_i \in V_1\otimes W_1 \\ &= \sum_{i\in\{1,\ldots,r\}} v_i\otimes y_i \in V_1\otimes W_2 \end{aligned} \tag{10.164}$$

in which r, k are minimal. Using 10.385 we have then that $\{u\}_{i\in\{1,\ldots,k\}}$, $\{v_i\}_{i\in\{1,\ldots,r\}}$, $\{x_i\}_{i\in\{1,\ldots,k\}}$, $\{y_i\}_{i\in\{1,\ldots,r\}}$ are linear independent sets. As $\{u_i\}_{i\in\{1,\ldots,k\}}$ is linear independent we have using 10.178 the existence of a linear mapping $h: V_1\rightarrow F$ such that $h(u_1)=1$ and $\forall i\in\{2,\ldots,k\}$ we have $h(u_i)=0$. Choose a arbitrary linear mapping $f: V_2\rightarrow F$ then we have that $h\cdot f: V_1\times V_2\rightarrow F$ is a multilinear mapping and thus by the definition of the tensor product there exists a unique linear $g: V_1\otimes V_2\rightarrow F$ such that

$$g\circ\otimes = h\cdot f \tag{10.165}$$

Using 10.164 we have then

$$\begin{aligned} g(x) &= g\left(\sum_{i\in\{1,\ldots,k\}} u_i\otimes x_i\right) \\ &= \sum_{i\in\{1,\ldots,k\}} g(u_i\otimes x_i) \\ &= \sum_{i\in\{1,\ldots,k\}} (g\circ\otimes)(u_i, x_i) \\ &= \sum_{i\in\{1,\ldots,k\}} h(u_i)\cdot f(x_i) \\ &= f(x_1) \end{aligned} \tag{10.166}$$

But we have also

$$\begin{aligned} g(x) &= g\left(\sum_{i\in\{1,\ldots,r\}} v_i\otimes y_i\right) \\ &= \sum_{i\in\{1,\ldots,r\}} g(v_i\otimes y_i) \\ &= \sum_{i\in\{1,\ldots,r\}} h(v_i)\cdot f(y_i) \\ &\underset{f is \text{linear}}{=} f\left(\sum_{i\in\{1,\ldots,r\}} h(v_i)\cdot y_i\right) \end{aligned}$$

so that we have for a arbitrary linear function $f: V_1 \rightarrow F$ that $f(x_1) = f\left(\sum_{i \in \{1,\ldots,r\}} h(v_i) \cdot y_i\right)$ which gives by 10.177 that $W_1 \ni x_1 = \sum_{i \in \{1,\ldots,r\}} h(v_i) \cdot y_i \in W_2$. This gives that $x_1 \in W_1 \bigcap W_2 = \{0\}$ or $x_1 = 0$ contradicting the linear Independence of $\{u_i\}_{i \in \{1,\ldots,k\}}$. So we conclude that our assumption of $x \neq 0$ is wrong and thus we must have that $(V_1 \otimes W_1) \bigcap (V_1 \otimes W_2) = \{0\}$. □

Lemma 10.393. *Let V_1, V_2 be vector spaces over a field F (with characterization zero) such that $V_1 = \sum_{i \in I}^{\oplus} W_i$ (or $V_2 = \sum_{i \in I}^{\oplus} W_i$) then we have that $V_1 \otimes V_2 = \sum_{i \in \{1,\ldots,m\}}^{\oplus} (W_i \otimes V_2)$ (or $V_1 \otimes V_2 = \sum_{i \in \{1,\ldots,m\}}^{\oplus} (V_1 \otimes W_i)$)*

Proof. Consider the two cases of the theorem

1. $\boldsymbol{(V_1 = \sum_{i \in I}^{\oplus} W_i)}$ We prove this by induction, so define $S = \left\{m \in \mathbb{N} \mid \text{if } W = \sum_{i \in \{1,\ldots,m\}}^{\oplus} W_i, V \text{ are vector spaces over } F \text{ then } W \otimes V = \sum_{i \in \{1,\ldots,m\}}^{\oplus} (W_i \otimes V)\right\}$ we have then that

   a. $\boldsymbol{(m = 1)}$ Here we have $W = \sum_{i \in \{1\}}^{\oplus} W_i = W_1$ (see 10.365) and thus $W \otimes V = W_1 \otimes V \underset{10.365}{=} \sum_{i \in \{1\}} (W_i \otimes V)$ proving that $1 \in S$

   b. $\boldsymbol{(m \in S)}$ Consider now $m + 1$ then using 10.368 we have that $W = \sum_{i \in \{1,\ldots,m+1\}}^{\oplus} W_i = \left(\sum_{i \in \{1,\ldots,\ldots,\}}^{\oplus} W_i\right) \oplus W_{m+1}$ and using the previous lemma we have then that $W \otimes V = \left(\left(\sum_{i \in \{1,\ldots,m\}}^{\oplus} W_i\right) \otimes V\right) \oplus (W_{m+1} \otimes V) \underset{m \in S}{=} (\sum_i^{\oplus} (W_i \otimes V)) \oplus (W_{m+1} \otimes V) \underset{10.368}{=} \sum_{i \in \{1,\ldots,m+1\}}^{\oplus} (W_i \otimes V)$ proving that $m + 1 \in S$

   Using induction we have then that $S = \mathbb{N}$ proving the first case (using $V_1 = W, V_2 = V$.

2. $\boldsymbol{(V_2 = \sum_{i \in I}^{\oplus} W_i)}$ We prove this by induction, so define $S = \left\{m \in \mathbb{N} \mid \text{if } V = \sum_{i \in \{1,\ldots,m\}}^{\oplus} W_i, W \text{ are vector spaces over } F \text{ then } W \otimes V = \sum_{i \in \{1,\ldots,m\}}^{\oplus} (W \otimes W_i)\right\}$ we have then that

   a. $\boldsymbol{(m = 1)}$ Here we have $V = \sum_{i \in \{1\}}^{\oplus} W_i = W_1$ (see 10.365) and thus $W \otimes V = W \otimes W_1 \underset{10.365}{=} \sum_{i \in \{1\}} (W \otimes W_i)$ proving that $1 \in S$

   b. $\boldsymbol{(m \in S)}$ Consider now $m + 1$ then using 10.368 we have that $V = \sum_{i \in \{1,\ldots,m+1\}}^{\oplus} W_i = \left(\sum_{i \in \{1,\ldots,\ldots,\}}^{\oplus} W_i\right) \oplus W_{m+1}$ and using the previous lemma we have then that $W \otimes V = \left(W \otimes \left(\sum_{i \in \{1,\ldots,m\}}^{\oplus} W_i\right)\right) \oplus (W \otimes W_{m+1}) \underset{m \in S}{=} (\sum_i^{\oplus} (W \otimes W_i)) \oplus (W \otimes W_{m+1}) \underset{10.368}{=} \sum_{i \in \{1,\ldots,m+1\}}^{\oplus} (W \otimes W_i)$ proving that $m + 1 \in S$

   Using induction we have then that $S = \mathbb{N}$ proving the second case (if $W = V_1$, $V = V_2$) □

Theorem 10.394. *Let V, W be vector spaces such that $V = \sum_{i \in \{1,\ldots,k\}}^{\oplus} W_i$, $W = \sum_{i \in \{1,\ldots,l\}}^{\oplus} U_i$. Then $V \otimes W = \sum_{\gamma \in \Gamma(k,l)}^{\oplus} W_{\gamma(1)} \otimes U_{\gamma(2)}$*

Proof. Again this is proved by induction so given $k\in\mathbb{N}$ and let $V=\sum^{\oplus}_{i\in\{1,\ldots,k\}} W_i$ let $S_k=\left\{m\in\mathbb{N}|\text{if } W=\sum^{\oplus}_{i\in\{1,\ldots,m\}} U_i \text{ then } V\otimes W=\sum^{\oplus}_{\gamma\in\Gamma(k,m)} W_{\gamma(1)}\otimes V_{\gamma(2)}\right\}$ then we have

1. **$(m=1)$** Then $W=\sum^{\oplus}_{i\in\{1\}} U_i=U_1$ and using the above lemma we have then that $V\otimes W=\sum^{\oplus}_{i\in\{1,\ldots,k\}} W_i\otimes U_1$, if we then consider the trivial bijection $\tau\colon\{1,\ldots,k\}\to\Gamma(k,1)$ by $\tau(i)(1)=i$ and $\tau(i)(2)=1$ then $V\otimes W=\sum^{\oplus}_{i\in\{1,\ldots,k\}} W_{\tau(i)(1)}\otimes U_{\tau(i)(2)} \underset{10.367}{=} \sum^{\oplus}_{\gamma\in\Gamma(k,1)} W_{\gamma(1)}\otimes U_{\gamma(2)}$ proving that $1\in S_k$

2. **$(m\in S)$** Then $W=\sum^{\oplus}_{i\in\{1,\ldots,m+1\}} U_i \underset{10.368}{=} \left(\sum^{\oplus}_{i\in\{1,\ldots,m\}} U_i\right)\oplus U_{m+1}$ so that by the above lemma we have $V\otimes W=\left(V\otimes\left(\sum^{\oplus}_{i\in\{1,\ldots,m\}} U_i\right)\right)\oplus (V\otimes U_{m+1})=\left(\left(\sum^{\oplus}_{i\in\{1,\ldots,k\}} W_i\right)\oplus\left(\sum^{\oplus}_{i\in\{1,\ldots,m\}} U_i\right)\right)\oplus\left(\left(\sum^{\oplus}_{i\in\{1,\ldots,k\}} W_i\right)\otimes U_{m+1}\right) \underset{m\in S \text{ and previous lemma}}{=} \left(\sum_{\gamma\in\Gamma(k,m)} W_{\gamma(1)}\otimes U_{\gamma(2)}\right)\oplus\left(\sum^{\oplus}_{i\in\{1\ldots,k\}} W_i\otimes U_{m+1}\right)$ proving that

$$V\otimes W=\left(\sum^{\oplus}_{\gamma\in\Gamma(k,m)} W_{\gamma(1)}\otimes U_{\gamma(2)}\right)\oplus\left(\sum^{\oplus}_{i\in\{1,\ldots,k\}} W_i\otimes U_{m+1}\right) \tag{10.167}$$

Define the trivial bijection $\kappa\colon\{1,\ldots,k\}\to\{1,\ldots,k\}\times\{m+1\}$ by $\kappa(i)(1)=i$ and $\kappa(i)(2)=m+1$ then we have $\sum^{\oplus}_{i\in\{1,\ldots,k\}} W_i\otimes U_{m+1}=\sum^{\oplus}_{i\in\{1,\ldots,k\}} W_{\kappa(i)(1)}\otimes U_{\kappa(i)(2)} \underset{10.367}{=} \sum_{\gamma\in\{1,\ldots,k\}\times\{m+1\}} W_{\gamma(1)}\otimes U_{\gamma(2)}$ giving

$$\sum^{\oplus}_{i\in\{1,\ldots,k\}} W_i\otimes U_{m+1}=\sum_{\gamma\in\{1,\ldots,k\}\times\{m+1\}} W_{\gamma(1)}\otimes U_{\gamma(2)} \tag{10.168}$$

As we have trivially that $\Gamma(k,m+1)=\{1,\ldots,k\}\times\{1,\ldots,m+1\}=\{1,\ldots,k\}\times\{1,\ldots,m\}\bigcup\{1,\ldots,k\}\times\{m+1\}=\Gamma(k,m)\bigcup\{1,\ldots,k\}\times\{m+1\}$ where the union is disjoint. Then using 10.167, 10.168 and 10.371 we have that $V\otimes W=\left(\sum^{\oplus}_{\gamma\in\Gamma(k,m)} W_{\gamma(1)}\otimes U_{\gamma(2)}\right)\oplus\sum_{\gamma\in\{1,\ldots,k\}\times\{m+1\}} W_{\gamma(1)}\otimes U_{\gamma(2)}=\sum_{\gamma\in\Gamma(k,m+1)} W_{\gamma(1)}\otimes U_{\gamma(2)}$. This proves that $m+1\in S$

Using induction we have then that $S_k=\mathbb{N}$. So as $l\in\mathbb{N}=S_k$ we have $V\otimes W=\sum^{\oplus}_{\gamma\in\Gamma(k,l)} W_{\gamma(1)}\otimes U_{\gamma(2)}$ □

Theorem 10.395. *Let $\{V_i\}_{i\in\{1,\ldots,m\}}$ be vector spaces over a field F such that $\forall i\in\{1,\ldots,m\}$ there exists a family $\{W_{i,j}\}_{j\in\{1,\ldots,k_i\}}$ of sub spaces of V_i such that $V_i=\sum^{\oplus}_{j\in\{1,\ldots,k_i\}} W_{i,j}$ (see $\oplus_{i\in\{1,\ldots,n\}}V_i$) then we have that $V_1\otimes\ldots\otimes V_n=\sum^{\oplus}_{\gamma\in\Gamma(k_1,\ldots,k_n)}(W_{1,\gamma(1)}\otimes\ldots\otimes W_{m,\gamma(m)})$ (where the $\otimes$ on the $W_{i,j}$ is defined by 10.384 based on $\otimes$ of the V_i's)*

Proof. We prove this by induction so let $S = \{m \in \mathbb{N}|\text{If } \{V_i\}_{i\in\{1,\ldots,m\}}$ be vector spaces over a filed F such that $\forall i \in \{1, \ldots, m\}$ there exists a family $\{W_{i,j}\}_{j\in\{1,\ldots,k_i\}}$ of subspaces of V_i such that $V_i = \sum_{j\in\{1,\ldots,k\}}^{\oplus} W_{i,j}$ then $V_1\otimes\ldots\otimes V_n = \sum_{\gamma\in\Gamma(k_1,\ldots,k_n)}^{\oplus}(W_{1,\gamma(1)}\otimes\ldots\otimes W_{m,\gamma(m)})\}$ then

1. **$(m=1)$** Then we have $V_1\otimes\ldots\otimes V_1 \underset{10.376}{=} V_1 = \sum_{j\in\{1,\ldots,k_1\}}^{\oplus} W_{1,j}$, using the trivial bijection κ: $\{1, \ldots, k_1\}\to\Gamma(k_1)$ by $\kappa(i)(1) = i$ then $\sum_{j\in\{1,\ldots,k_j\}}^{\oplus} W_{1,\kappa(j)(1)} = \sum_{\gamma\in\Gamma(k_1)} W_{1,\gamma(1)} = \sum_{\gamma\in\Gamma(k_1)} W_{1,\gamma(1)}\otimes\ldots\otimes W_{1,\gamma(1)}$ proving that $1\in S$

2. **$(m\in S)$** Then we have for $m+1$ that $V_1\otimes\ldots\otimes V_{m+1} \underset{10.386}{=} (V_1\otimes\ldots\otimes V_m)\otimes V_{m+1}$ (we can find a tensor product of $V_1\otimes\ldots\otimes V_m$ and V_{m+1} satisfying the equation). $V_1\otimes\ldots\otimes V_m \underset{m\in S}{=} \sum_{\gamma\in\Gamma(k_1,\ldots,k_m)}^{\oplus} W_{1,\gamma(1)}\otimes\ldots\otimes W_{m,\gamma(m)}$ and thus

$$\begin{aligned}
(V_1\otimes\ldots\otimes V_m)\otimes V_{m+1} &= \left(\sum_{\gamma\in\Gamma(k_1,\ldots,k_m)}^{\oplus} W_{1,\gamma(1)}\otimes\ldots\otimes W_{m,\gamma(m)}\right)\otimes V_{m+1} \\
&= \left(\sum_{\gamma\in\Gamma(k_1,\ldots,k_m)}^{\oplus} ((W_{1,\gamma(1)}\otimes\ldots\otimes W_{m,\gamma(m)}))\right)\otimes \\
&\quad \sum_{\iota\in\{1,\ldots,k_{m+1}\}}^{\oplus} W_{m+1,i} \\
&\underset{10.393}{=} \sum_{i\in\{1,\ldots,k_{m+1}\}}^{\oplus}\left(\left(\sum_{\gamma\in\Gamma(k_1,\ldots,k_m)}^{\oplus} ((W_{1,\gamma(1)}\otimes\ldots\otimes \right.\right. \\
&\quad \left.\left. W_{m,\gamma(m)}))\right)\otimes W_{m+1,i}\right) \\
&\underset{10.393}{=} \sum_{i\in\{1,\ldots,k_{m+1}\}}^{\oplus}\left(\sum_{\gamma\in\Gamma(k_1,\ldots,k_m)}^{\oplus} (W_{1,\gamma(1)}\otimes\ldots\otimes \right. \\
&\quad \left. W_{m,\gamma(m)})\otimes W_{m+1,i}\right) \\
&\underset{10.386}{=} \sum_{i\in\{1,\ldots,k_{m+1}\}}^{\oplus}\left(\sum_{\gamma\in\Gamma(k_1,\ldots,k_m)}^{\otimes} W_{1,\gamma(1)}\otimes\ldots\otimes \right. \\
&\quad \left. W_{m,\gamma(m)}\otimes W_{m+1,i}\right)
\end{aligned} \tag{10.169}$$

If we define now $\{1, \ldots, k_{m+1}\} \times \Gamma(k_1, \ldots, k_m) \underset{\text{disjoint union}}{=} \bigcup_{i \in \{1, \ldots, k_m\}} \{i\} \times \Gamma(k_1, \ldots, k_m)$ then 10.169 becomes

$$(V_1 \otimes \ldots \otimes V_m) \otimes V_{m+1} \underset{10.372}{=} \sum_{(i,\gamma) \in \{1, \ldots, k_{m+1}\} \times \Gamma(k_1, \ldots, k_m)} W_{1,\gamma(1)} \otimes \ldots \otimes W_{m,\gamma(m)} \otimes W_{m+1,i} \tag{10.170}$$

Define the function $\chi: \{1, \ldots, k_{m+1}\} \times \Gamma(k_1, \ldots, k_m) \rightarrow \Gamma(k_1, \ldots, k_{m+1})$ where $\chi(i, \gamma)(j) = \begin{cases} \gamma(j) \text{ if } j \leqslant m \\ i \text{ if } j = m+1 \end{cases}$ then we have

a. **(injectivity)** If $\chi(i, \gamma) = \chi(i', \gamma')$ then we have $\chi(i, \gamma)(m+1) = \chi(i', \gamma')(m+1) \Rightarrow i = i'$, also if $j \in \{1, \ldots, m\}$ then $\chi(i, \gamma)(j) = \chi(i', \gamma)(j) \Rightarrow \gamma(j) = \gamma'(j) \Rightarrow \gamma = \gamma'$ so that $(i, \gamma) = (i', \gamma')$

b. **(surjectivity)** If $\gamma \in \Gamma(k_1, \ldots, k_{m+1})$ construct then $(\gamma(m+1), \gamma_{|\{1, \ldots, m\}})$ and then we have if $j \in \{1, \ldots, m+1\}$ that $\chi(\gamma(m+1)) = \begin{cases} \gamma_{|\{1, \ldots, m\}}(j) \text{ if } j \leqslant m \\ \gamma(m+1) \text{ if } j = m+1 \end{cases} = \begin{cases} \gamma(j) \text{ if } j \leqslant m \\ \gamma(m+1) \text{ if } j = m+1 \end{cases} = \gamma(j) \Rightarrow \mathcal{X}(\gamma(m+1), \gamma_{|\{1, \ldots, m\}}) = \gamma$

so χ is a bijection and we can rewrite 10.170 as follows

$$\begin{aligned} V_1 \otimes \ldots \otimes V_{m+1} &= (V_1 \otimes \ldots \otimes V_m) \otimes V_{m+1} \\ &= \sum_{(i,\gamma) \in \{1, \ldots, k_{m+1}\} \times \Gamma(k_1, \ldots, k_m)} W_{1,\chi(1,\gamma)(1)} \otimes \ldots \otimes W_{1,\chi(m,\gamma)(m)} \otimes W_{m+1,\chi(m+1,\gamma)(m+1)} \\ &\underset{10.367}{=} \sum_{\gamma \in \Gamma(k_1, \ldots, k_{m+1})} W_{1,\gamma} \otimes \ldots \otimes W_{m+1,\gamma(m+1)} \end{aligned}$$

This proves that $m+1 \in S$

By induction we have then that $S = \mathbb{N}$ proving the theorem □

We use now the above theorem to construct the basis of a tensor product out of the basis of the factors.

First we prove the following lemma

Lemma 10.396. *Let $\{V_i\}_{i \in \{1, \ldots, m\}}$ be a family of one dimensional vector spaces ($\forall i \in \{1, \ldots, m\}$ we have that $\exists e_i \in V_i, e_i \neq 0$ such that $V_i = \mathcal{S}(\{e_i\})$) then every tensor product $V_1 \otimes \ldots \otimes V_m = \mathcal{S}(\{e_1 \otimes \ldots \otimes e_n\})$*

Proof. First we prove that $\langle P, \nu \rangle$ where $P = \mathcal{S}(\{e_1 \otimes \ldots \otimes e_m\})$ and $\nu: \prod_{i \in \{1, \ldots, m\}} \mathcal{S}(\{e_i\}) \rightarrow \mathcal{S}(\{e_1 \otimes \ldots \otimes e_m\})$ is defined by $\nu(\lambda_1 \cdot e_1, \ldots, \lambda_n \cdot e_n) = \left(\prod_{i \in \{1, \ldots m\}} \lambda_i\right) \cdot (e_1 \otimes \ldots \otimes e_m)$ is a tensor product of $\{V_i\}_{i \in \{1, \ldots, m\}}$

1. P is a vector space by 10.130

2. If $\alpha, \beta \in F$ then $\nu(\lambda_1 \cdot e_i, \ldots, \lambda_{i-1} \cdot e_{i-1}, \alpha \cdot (\lambda'_i \cdot e_i) + \beta \cdot (\lambda''_i \cdot e_i), \lambda_{i+1} \cdot e_{i+1}, \ldots, \lambda_m \cdot e_m) = \nu(\lambda_1 \cdot e_1, \ldots, \lambda_{i-1} \cdot e_{i-1}, (\alpha \cdot \lambda'_i + \beta \cdot \lambda''_i) \cdot e_i, \lambda_{i+1} \cdot e_{i+1}, \ldots, \lambda_m \cdot e_m) = \left(\prod_{i \in \{1,\ldots,m\}} \lambda_i\right) \cdot (e_1 \otimes \ldots \otimes e_m)$ where $\{\lambda_i\}_{i \in \{1,\ldots,\}}$ is defined by $\lambda_j = \begin{cases} \lambda_j \text{ if } j \neq i \\ \alpha \cdot \lambda'_i + \beta \cdot \lambda''_i \text{ if } j = i \end{cases}$ and then by 10.230 we have that $\left(\prod_{i \in \{1,\ldots,m\}} \lambda_i\right) = \alpha \cdot \left(\prod_{i \in \{1,\ldots,m\}} (\lambda_1, \ldots, \lambda_{i-1}, \lambda'_i, \lambda_{i+1}, \ldots, \lambda_m)_i\right) + \beta \cdot \left(\prod_{i \in \{1,\ldots,m\}} (\lambda_1, \ldots, \lambda_{i-1}, \lambda''_i, \lambda_{i+1}, \ldots, \lambda_m)_i\right) \Rightarrow \nu(\lambda_1 \cdot e_i, \ldots, \lambda_{i-1} \cdot e_{i-1}, \alpha \cdot (\lambda'_i \cdot e_i) + \beta \cdot (\lambda''_i \cdot e_i), \lambda_{i+1} \cdot e_{i+1}, \ldots, \lambda_m \cdot e_m) = \alpha \cdot \left(\prod_{i \in \{1,\ldots,m\}} (\lambda_1, \ldots, \lambda_{i-1}, \lambda'_i, \lambda_{i+1}, \ldots, \lambda_m)_i\right) \cdot (e_1 \otimes \ldots \otimes e_m) + \beta \cdot \left(\prod_{i \in \{1,\ldots,m\}} (\lambda_1, \ldots, \lambda_{i-1}, \lambda''_i, \lambda_{i+1}, \ldots, \lambda_m)_i\right) \cdot (e_1 \otimes \ldots \otimes e_m) = \alpha \cdot \nu(\lambda_1 \cdot e_1, \ldots, \lambda_{i-1} \cdot e_{i-1}, \lambda'_i \cdot e_i, \lambda_{i+1} \cdot e_{i+1}, \ldots, \lambda_m \cdot e_m) + \beta \cdot \nu(\lambda_1 \cdot e_1, \ldots, \lambda_{i-1} \cdot e_{i-1}, \lambda''_i \cdot e_i, \lambda_{i+1} \cdot e_{i+1}, \ldots, \lambda_m \cdot e_m)$ proving that ν is multilinear$\Rightarrow \nu \in \operatorname{Hom}(\mathcal{S}(\{e_1\}), \ldots, \mathcal{S}(\{e_m\}), \mathcal{S}(\{e_1 \otimes \ldots \otimes e_m\}))$

3. As by definition of ν we have $\nu\left(\prod_{i \in \{1,\ldots,m\}} \mathcal{S}(\{e_i\})\right) \subseteq \mathcal{S}(\{e_1 \otimes \ldots \otimes e_m\})$ we have $\mathcal{S}\left(\nu\left(\prod_{i \in \{1,\ldots,m\}} \mathcal{S}(\{e_i\})\right)\right) \subseteq \mathcal{S}(\mathcal{S}(\{e_1 \otimes \ldots \otimes e_m\})) = \mathcal{S}(\{e_1 \otimes \ldots \otimes e_m\}) = P \Rightarrow \mathcal{S}\left(\nu\left(\prod_{i \in \{1,\ldots,m\}} \mathcal{S}(\{e_i\})\right)\right) \subseteq P$. If $v \in P = \mathcal{S}(\{e_1 \otimes \ldots \otimes e_m\}) \Rightarrow v = \lambda \cdot (e_1 \otimes \ldots \otimes e_m)$ and $\nu(\lambda \cdot e_1, 1 \cdot e_2, \ldots, 1 \cdot e_m) = v \Rightarrow P \subseteq \nu\left(\prod_{i \in \{1,\ldots,m\}} \mathcal{S}(\{e_i\})\right) \Rightarrow P \underset{P \text{ is a vactor space}}{=} \mathcal{S}(P) \subseteq \mathcal{S}\left(\nu\left(\prod_{i \in \{1,\ldots,m\}} \mathcal{S}(\{e_i\})\right)\right)$ proving that $P = \mathcal{S}\left(\nu\left(\prod_{i \in \{1,\ldots,m\}} \mathcal{S}(\{e_i\})\right)\right)$

4. If $\varphi \in \operatorname{Hom}(\mathcal{S}(\{e_1\}), \ldots, \mathcal{S}(\{e_m\}), U)$ define then $h: \mathcal{S}(\{e_1 \otimes \ldots \otimes e_m\}) \rightarrow U$ by $\lambda \cdot (e_1 \otimes \ldots \otimes e_m) \rightarrow h(\lambda \cdot (e_1 \otimes \ldots \otimes e_m)) = \lambda \cdot \varphi(e_1, \ldots, e_m)$ which is trivially linear. Then we have $\varphi(\lambda_1 \cdot e_1, \ldots, \lambda_m \cdot e_m) \underset{10.239}{=} \left(\prod_{i \in \{1,\ldots,m\}} \lambda_i\right) \cdot \varphi(e_1, \ldots, e_m) = h\left(\left(\prod_{i \in \{1,\ldots,m\}}\right) \cdot (e_1 \otimes \ldots \otimes e_m)\right) = h(\nu(\lambda_1 \cdot e_1, \ldots, \lambda_m \cdot e_m))$ proving that $h \circ \nu = \varphi$.

So we have proved that $\langle P, \nu \rangle$ is a tensor product of $\{V_i\}_{i \in \{1,\ldots,m\}}$ which is obviously one dimensional as by 10.377 $e_1 \otimes \ldots \otimes e_m = \nu(e_1, \ldots, e_m) \neq 0$. If now $\langle Q, \mu \rangle$ is another tensor product of $\{V_i\}_{i \in \{1,\ldots,m\}}$ then by 10.374 we have that Q is isomorphic with P, using 10.220 we have then that Q is also one dimensional. So there exist a $v_0 \in Q$ such that $Q = \mathcal{S}(\{v_0\})$. Now as by 10.377 e_1$\otimes \ldots \otimes e_m = \mu(e_1, \ldots, e_m) \neq 0$ we have

1. $x \in \mathcal{S}(e_1 \otimes \ldots \otimes e_m) \Rightarrow x = \lambda \cdot (e_1 \otimes \ldots \otimes e_m) \Rightarrow x \in Q$ so $\mathcal{S}(e_1 \otimes \ldots \otimes e_m) \subseteq Q$
2. if $x \in Q$ then $x = \lambda \cdot v_0$, now as $0 \neq e_1 \otimes \ldots \otimes e_m$ there exists a $\tau \neq 0$ such that $e_1 \otimes \ldots \otimes e_m = \tau \cdot v_0 \Rightarrow v_0 = \frac{1}{\tau} \cdot e_1 \otimes \ldots \otimes e_m$ so that $x = \left(\lambda \cdot \frac{1}{\tau}\right) \cdot e_1 \otimes \ldots \otimes e_m$ proving that $Q \subseteq \mathcal{S}(\{e_1 \otimes \ldots \otimes e_m\})$

(1) and (2) gives finally that $Q = \mathcal{S}(\{e_1 \otimes \ldots \otimes e_m\})$ for all tensor spaces □

Theorem 10.397. *Let $\{V_i\}_{i \in \{1,\ldots,m\}}$ be finite dimensional vector spaces over a field F with $\forall i \in I$ $\{e_{i,j}\}_{j \in \{1,\ldots,k_i\}}$ a basis of V_i then $\{e_{1,\gamma(1)} \otimes \ldots \otimes e_{m,\gamma(m)}\}_{\gamma \in \Gamma(k_1,\ldots,k_m)}$ is a basis of $V_1 \otimes \ldots \otimes V_m$. So the dimension of $V_1 \otimes \ldots \otimes V_m$ is $\prod_{i \in \{1,\ldots,m\}} k_i$*

Proof. First remark that from 10.151 and the definition of a direct sum if follows that $\forall i \in I, \forall j \in \{1, \ldots, k_i\}$ we have that $\mathcal{S}(\{e_{i,j}\}) = \{\lambda \cdot e_{i,j} | \lambda \in F\}$ and using 10.151 we have then that $V_i = \sum_{j \in \{1,\ldots,k_i\}}^{\oplus} \mathcal{S}(e_{i,j})$. Using 10.395 we have then that $V_1 \otimes \ldots \otimes V_m = \sum^{\otimes}_{\gamma \in \Gamma(k_1,\ldots,k_m)} (\mathcal{S}(\{e_{1,\gamma(1)}\}) \otimes \ldots \otimes \mathcal{S}(\{e_{m,\gamma(m)}\}))$. As $\forall i \in I$, $\forall j \in \{1,\ldots,k_i\}$ we have $e_{i,j} \neq 0$ (they are part of a basis) we use the previous lemma to prove that $\mathcal{S}(\{e_{1,\gamma(1)}\}) \otimes \ldots \otimes \mathcal{S}(\{e_{m,\gamma(m)}\}) = \mathcal{S}(e_{1,\gamma(1)} \otimes \ldots \otimes e_{m,\gamma(m)})$. This gives $V_1 \otimes \ldots \otimes V_m = \sum^{\otimes}_{\gamma \in \Gamma(k_1,\ldots,k_m)} \mathcal{S}(e_{1,\gamma(1)} \otimes \ldots \otimes e_{m,\gamma(m)})$ proving by 10.151 and the definition of a direct sum that $\{e_{1,\gamma(1)} \otimes \ldots \otimes e_{m,\gamma(m)}\}_{\gamma \in \Gamma(k_1,\ldots,k_m)}$ forms a basis of $V_1 \otimes \ldots \otimes V_m$. □

Using the above theorem we can then define the components of a the tensor product of finite dimensional spaces.

Definition 10.398. *Let $\{V_i\}_{i \in \{1,\ldots,m\}}$ be finite dimensional vector spaces over a field F with $\forall i \in I$ $\{e_{i,j}\}_{j \in \{1,\ldots,k_i\}}$ a basis of V_i then by the previous theorem we have that $\{e_{1,\gamma(1)} \otimes \ldots \otimes e_{m,\gamma(m)}\}_{\gamma \in \Gamma(k_1,\ldots,k_m)}$ is a basis of $V_1 \otimes \ldots \otimes V_m$ so that we can find for a $T \in V_1 \otimes \ldots \otimes V_m$ a $\{T_\gamma\}_{\gamma \in \Gamma(k_1,\ldots,k_m)}$ so that $T = \sum_{\gamma \in \Gamma(k_1,\ldots,k_m)} T_\gamma \cdot (e_{1,\gamma(1)} \otimes \ldots \otimes e_{m,\gamma(m)})$ the family $\{T_\gamma\}_{\gamma \in (k_1,\ldots,k_m)}$ of $\prod_{i \in 1} k_i$ elements of F is called the family of components of the tensor T. As a element $\gamma \in \Gamma(k_1,\ldots,k_m)$ is written as $(\gamma(1),\ldots,\gamma(m)) = (\gamma_1,\ldots,\gamma_m)$ where $\forall i \in \{1,\ldots,m\}$ we have $\gamma(i) = \gamma_i \in \{1,\ldots,k_i\}$ we can write the components also as $\{T_{\gamma_1,\ldots,\gamma_m}\}_{\gamma \in \Gamma(k_1,\ldots,k_m)}$*

Definition 10.399. *If V is a vector space over a field F then if $p,q \in \mathbb{N}$ then $V^p_q = V_1 \otimes \ldots. \otimes V_m$ where $m = q + p$ and $\forall i \in \{1,\ldots,p\}$ we have $V_i = V$ and $\forall i \in \{p+1,\ldots, p+q=m\}$ $V_i = V^*$ (see 10.186). Elements of V^p_q are called p-contra-variant,q-co-variant tensors. As a shorthand we write $V^p_q = \overbrace{V \otimes \ldots \otimes V}^{p} \otimes \overbrace{V^* \otimes \ldots \otimes V^*}^{q}$. If $p,q = 0$ then we define $V^0_0 = F$*

Theorem 10.400. *Let V be a vector space over a field F then if $p,q \in \mathbb{N}$. For $s \in \{1,\ldots,p\}$, $t \in \{p+1,\ldots,p+q\}$ define $\varphi_{s,t}$: $\overbrace{V \times \ldots \times V}^{p} \times \overbrace{V^* \times \ldots \times V^*}^{q} \to V^{p-1}_{q-1}$ defined by $\varphi_{s,t}(v_1,\ldots,v_p,v_{p+1},\ldots,v_{p+q}) = v_t(v_s) \cdot v_1 \otimes \ldots \otimes v_{s-1} \otimes v_{s+1} \otimes \ldots \otimes v_p \otimes v_{p+1} \otimes \ldots \otimes v_{t-1} \otimes v_{t+1} \otimes \ldots \otimes v_{p+q}$ then $\varphi_{s,t}$ is multilinear and thus $\varphi_{s,t} \in \mathrm{Hom}\left(\overbrace{V,\ldots,V}^{p}, \overbrace{V^*,\ldots,V^*}^{q}, V^{p-1}_{a-1}\right)$. By definition of the tensor product there exists a mapping $c_{s,t}$: $V^p_q \to V^{p-1}_{q-1}$ such that $c_{s,t} \circ \otimes = \varphi_{s,t}$ or $c_{s,t}(v_1 \otimes \ldots \otimes v_p \otimes v_{p+1} \otimes \ldots \otimes v_{p+q}) = v_t(v_s) \cdot v_1 \otimes \ldots \otimes v_{s-1} \otimes v_{s+1} \otimes \ldots \otimes v_p \otimes v_{p+1} \otimes \ldots \otimes v_{t-1} \otimes v_{t+1} \otimes \ldots \otimes v_{p+q}$. $c_{s,t}$ is called a contraction mapping*

Proof. To prove that $\varphi_{s,t}$ is linear take $i \in \{1,\ldots,p+q\}$ then we have either

1. $\boldsymbol{(i = s)}$ then $\varphi_{s,t}(v_1,\ldots,v_{i-1},v_{i+1},\ldots,v_{p+q}) = \varphi_{s,t}(v_1,\ldots v_{s-1}, \alpha \cdot v + \beta \cdot v', v_{s+1},\ldots,v_p,v_{p+1},\ldots,v_{p+q}) = v_t(\alpha \cdot v + \beta \cdot v') \cdot v_1 \otimes \ldots \otimes v_{s-1} \otimes v_{s+1} \otimes \ldots \otimes v_p \otimes v_{p+1} \otimes \ldots \otimes v_{t-1} \otimes v_{t+1} \otimes \ldots \otimes v_{p+q} \underset{v_t \in V^*}{=} \alpha \cdot v_t(v) \cdot v_1 \otimes \ldots \otimes v_{s-1} \otimes v_{s+1} \otimes \ldots \otimes v_p \otimes v_{p+1} \otimes \ldots \otimes v_{t-1} \otimes v_{t+1} \otimes \ldots \otimes v_{p+q} + \beta \cdot v_t(v') \cdot v\_1 \otimes \ldots \otimes v\_(s-1) \otimes v\_(s+1) \otimes \ldots \otimes v\_p \otimes v\_(p+1) \otimes \ldots \otimes v\_(t-1) \otimes v\_(t+1) \otimes \ldots \otimes v\_(p+q) = \alpha \cdot \varphi_{s,t}(v_1,\ldots v_{s-1},v,v_{s+1},\ldots,v_p,v_{p+1},\ldots,v_{p+q}) + \beta \cdot \varphi_{s,t}(v_1,\ldots v_{s-1},v',v_{s+1},\ldots,v_p,v_{p+1},\ldots,v_{p+q}) = \alpha \cdot \varphi_{s,t}(v_1,\ldots,v_{i-1},v,v_{i+1},\ldots,v_{p+q}) + \beta \cdot \varphi_{s,t}(v_1,\ldots,v_{i-1},v',v_{i+1},\ldots,v_{p+q})$

2. $\boldsymbol{(i=t)}$ then $\varphi_{s,t}(v_1, ...v_{i-1}, \alpha\cdot v+\beta\cdot v', v_{i+1}, ..., v_{p+q}) = \varphi_{s,t}(v_1, ...v_p, v_{p+1}, ...v_{t-1}, \alpha\cdot v+\beta\cdot v', v_{t+1}, ..., v_{p+q}) = (\alpha\cdot v+\beta\cdot v')(v_s)\cdot v_1\otimes ...\otimes v_{s-1}\otimes v_{s+1}\otimes ...\otimes v_p\otimes v_{p+1}\otimes ...\otimes v_{t-1}\otimes v_{t+1}\otimes ...\otimes v_{p+q} = \alpha\cdot v(v_s)\cdot v_1\otimes ...\otimes v_{s-1}\otimes v_{s+1}\otimes ...\otimes v_p\otimes v_{p+1}\otimes ...\otimes v_{t-1}\otimes v_{t+1}\otimes ...\otimes v_{p+q}+\beta\cdot v'(v_s)\cdot v_1\otimes ...\otimes v_{s-1}\otimes v_{s+1}\otimes ...\otimes v_p\otimes v_{p+1}\otimes ...\otimes v_{t-1}\otimes v_{t+1}\otimes ...\otimes v_{p+q} = \alpha\cdot\varphi_{s,t}(v_1, ...v_p, v_{p+1}, ...v_{t-1}, v, v_{t+1}, ..., v_{p+q})+\beta\cdot\varphi_{s,t}(v_1, ...v_p, v_{p+1}, ...v_{t-1}, v', v_{t+1}, ..., v_{p+q}) = \alpha\cdot\varphi_{s,t}(v_1, ..., v_{i-1}, v, v_{i+1}, ..., v_{p+q})+\beta\cdot\varphi_{s,t}(v_1, ..., v_{i-1}, v', v_{i+1}, ..., v_{p+q})$

3. $\boldsymbol{(i\neq s,t)}$ then we have the following sub-cases to consider as $s<t$

   a. $\boldsymbol{(i<s<t)}$ $\varphi_{s,t}(v_1, ...v_{i-1}, \alpha\cdot v+\beta\cdot v', v_{i+1}, ..., v_{p+q}) = v_t(v_s)\cdot v_1\otimes ...\otimes v_{i-1}\otimes(\alpha\cdot v+\beta\cdot v')\otimes v_{i+1}\otimes ...\otimes v_{s-1}\otimes v_{s+1}\otimes ...\otimes v_{t-1}\otimes v_{t+1}\otimes ...\otimes v_{p+q} \underset{\otimes\,\text{is multilinear}}{=} \alpha\cdot v_t(v_s)\cdot v_1\otimes ...\otimes v_{i-1}\otimes v\otimes v_{i+1}\otimes ...\otimes v_{s-1}\otimes v_{s+1}\otimes ...\otimes v_{t-1}\otimes v_{t+1}\otimes ...\otimes v_{p+q}+\beta\cdot v_t(v_s)\cdot v_1\otimes ...\otimes v_{i-1}\otimes v'\otimes v_{i+1}\otimes ...\otimes v_{s-1}\otimes v_{s+1}\otimes ...\otimes v_{t-1}\otimes v_{t+1}\otimes ...\otimes v_{p+q} = \alpha\cdot\varphi_{s,t}(v_1, ..., v_{i-1}, v, v_{i+1}, ..., v_{p+q})+\beta\cdot\varphi_{s,t}(v_1, ..., v_{i-1}, v', v_{i+1}, ..., v_{p+q})$

   b. $\boldsymbol{(s<i<t)}$ $\varphi_{s,t}(v_1, ..., v_{i-1}, \alpha\cdot v+\beta\cdot v', v_{i+1}, ..., v_{p+q}) = v_t(v_s)\cdot v_1\otimes ...\otimes v_{s-1}\otimes v_{s+1}\otimes ...\otimes v_{i-1}\otimes(\alpha\cdot v+\beta\cdot v')\otimes ...\otimes v_{t-1}\otimes v_{t+1}\otimes ...\otimes v_{p+q} \underset{\otimes\,\text{is multilinear}}{=} \alpha\cdot v_t(v_s)\cdot v_1\otimes ...\otimes v_{s-1}\otimes v_{s+1}\otimes ...\otimes v_{i-1}\otimes v\otimes ...\otimes v_{t-1}\otimes v_{t+1}\otimes ...\otimes v_{p+q}+\beta\cdot v_t(v_s)\cdot v_1\otimes ...\otimes v_{s-1}\otimes v_{s+1}\otimes ...\otimes v_{i-1}\otimes v'\otimes ...\otimes v_{t-1}\otimes v_{t+1}\otimes ...\otimes v_{p+q} = \alpha\cdot\varphi_{s,t}(v_1, ..., v_{i-1}, v, v_{i+1}, ..., v_{p+q})+\beta\cdot\varphi_{s,t}(v_1, ..., v_{i-1}, v', v_{i+1}, ..., v_{p+q})$

   c. $\boldsymbol{(s<t<i)}$ $\varphi_{s,t}(v_1, ..., v_{i-1}, \alpha\cdot v+\beta\cdot v', v_{i+1}, ..., v_{p+q}) = v_t(v_s)\cdot v_1\otimes ...\otimes v_{s-1}\otimes v_{s+1}\otimes ...\otimes v_{t-1}\otimes v_{t+1}\otimes ...\otimes v_{i-1}\otimes(\alpha\cdot v+\beta\cdot v')\otimes v_{i+1}\otimes ...\otimes v_{p+q} \underset{\otimes\,\text{is multilinear}}{=} \alpha\cdot v_t(v_s)\cdot v_1\otimes ...\otimes v_{s-1}\otimes v_{s+1}\otimes ...\otimes v_{t-1}\otimes v_{t+1}\otimes ...\otimes v_{i-1}\otimes v\otimes v_{i+1}\otimes ...\otimes v_{p+q}+\beta\cdot v_t(v_s)\cdot v_1\otimes ...\otimes v_{s-1}\otimes v_{s+1}\otimes ...\otimes v_{t-1}\otimes v_{t+1}\otimes ...\otimes v_{i-1}\otimes v'\otimes v_{i+1}\otimes ...\otimes v_{p+q} = \alpha\cdot\varphi_{s,t}(v_1, ..., v_{i-1}, v, v_{i+1}, ..., v_{p+q})+\beta\cdot\varphi_{s,t}(v_1, ..., v_{i-1}, v', v_{i+1}, ..., v_{p+q})$

(1),(2) and (3) proves then that $\varphi_{s,t}$ is multilinear. □

Chapter 12

Topology

12.1 Topological spaces

Definition 12.1. *A topological space $\langle X, \mathcal{T}\rangle$ is a set X together with $\mathcal{T} \subseteq \mathcal{P}(X)$ (a set of subsets of X) such that:*

1. $\emptyset \in \mathcal{T}$
2. $X \in \mathcal{T}$
3. $\forall U, B \in \mathcal{T}$ *we have* $U \bigcap V \in \mathcal{T}$
4. *If* $\{U_i\}_{i\in I}$ *is a family of sets in* $\mathcal{T}$ *then* $\bigcup_{i\in I} U_i \in \mathcal{T}$

We call $\mathcal{T}$ the topology of X and elements of $\mathcal{T}$ open sets.

Example 12.2. $\langle \emptyset, \{\emptyset\}\rangle$ is a trivially a topological space

Example 12.3. $\langle X, \mathcal{P}(X)\rangle$ is a trivially a topological space.

Theorem 12.4. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and let $\{U_i\}_{i\in I}$ be a finite family of open sets then $\bigcap_{i\in I} U_i$ is open.*

Proof. As I is finite we have the following possibilities

1. **($I = \emptyset$)** then if $x \in X$ we have $\forall i \in I | x \in U_i$ is fullfilled vacuously so $\bigcap_{i\in I} U_i = X \in \mathcal{T}$
2. **($I \neq \emptyset$)** then there exists a $n \in \mathbb{N}$ and a bijection $i: \{1, \ldots, n\} \to I$ so that $\bigcap_{k\in I} U_k = \bigcap_{j\in\{1,\ldots,n\}} U_{i(j)}$. We prove the rest by induction so let $X = \left\{n \in \mathbb{N} | \text{if } \{U_{i(j)}\}_{j\in\{1,\ldots,n\}} \text{ is open then } \bigcap_{j\in\{1,\ldots,n\}} U_{i(j)} \text{ is open}\right\}$ then:
   a. If $n = 1$ then $\bigcap_{j\in\{1\}} U_{i(j)} = U_{i(1)} \in \mathcal{T}$
   b. If $n \in X$ then $\bigcup_{j\in\{1,\ldots,n+1\}} U_{i(j)} \underset{1.107}{=} \left(\bigcup_{j\in\{1,\ldots,n\}} U_{i(j)}\right) \bigcup U_{i(j+1)}$ which is open because of $n \in X$ and (3) of the definition of a topology. □

Theorem 12.5. *Let $\langle X, \mathcal{T}\rangle$ be a topological space then we have $U \in \mathcal{T}$ iff $\forall x \in U$ we have $\exists V_x \in \mathcal{T}$ such that $x \in V_x \subseteq U$*

Proof.

1. **($\Rightarrow$)** If $U \in \mathcal{T}$ the we have either
   a. **($U = \emptyset$)** then $\forall x \in \emptyset \vDash \exists V_x \in \mathcal{T} \vdash x \in V_x \subseteq \emptyset$ is satisfied vacuously
   b. **($U \neq \emptyset$)** then $\forall x \in U$ we have $x \in U \subseteq U$
2. **($\Leftarrow$)** By the hypothesis we have that $\forall x \in U$ the set $\mathcal{A}_x = \{V \in \mathcal{T} | x \in V \subseteq U\}$ is not empty, so by the axiom of choice (see 2.201) that there exists a function $j: U \rightarrow \bigcup_{x \in U} \mathcal{A}_x$ and this defines then a family $\{j(x)\}_{x \in U}$ of open sets such that $x \in j(x) \subseteq U$ so that $U = \bigcup_{x \in U} j(x)$ is open by the definition of a topology. $\square$

We define now the subspace topology

Theorem 12.6. *Let $\langle X, \mathcal{T} \rangle$ be a topological space and $Y \subseteq X$ then if we define $\mathcal{T}_Y = \{U \bigcap Y | U \in \mathcal{T}\}$ we have that $\langle Y, \mathcal{T}_Y \rangle$ forms a topological space. $\mathcal{T}_Y$ is called the subspace topology induced by $\mathcal{T}$ on Y.*

Proof.

1. $\emptyset = \emptyset \bigcap Y \in \mathcal{T}_Y$
2. $Y = X \bigcap Y \in \mathcal{T}_Y$
3. If $U_1, U_2 \in \mathcal{T}_Y$ then there exists $V_1, V_2 \in \mathcal{T}$ such that $U_1 = V_1 \bigcap Y$, $U_2 = V_2 \bigcap Y$ then $U_1 \bigcap U_2 = (V_1 \bigcap Y) \bigcap (V_2 \bigcap Y) = (V_1 \bigcap V_2) \bigcap (Y \bigcap Y) = (V_1 \bigcap V_2) \bigcap Y \in \mathcal{T}_Y$ as $V_1 \bigcap V_2 \in \mathcal{T}$
4. If $\{U_i\}_{i \in I}$ is a family of open sets in $\mathcal{T}$ then $\forall i \in I$ there exists a V_i such that $U_i = V_i \bigcap Y$, this defines a family $\{V_i\}_{i \in I}$ then $\bigcup_{i \in I} U_i = \bigcup_{i \in I} (V_i \bigcap Y) \underset{1.107}{=} (\bigcup_{i \in I} V_i) \bigcap Y \in \mathcal{T}_Y$ as $\bigcup_{i \in I} V_i \in \mathcal{T}$ $\square$

Remark 12.7. If $\langle X, \mathcal{T} \rangle$ is a topological space and $U \in \mathcal{T}$ then if $V \in \mathcal{T}_U$ then $\exists V' \in \mathcal{T}$ such that $V = V' \bigcap U \in \mathcal{T}$. So in the case of $U \in \mathcal{T}$ (or U is open) we have $\mathcal{T}_U \subseteq \mathcal{T}$ (all open sets in the subspace topology induced on a open set are open in the inducing topology).

The following theorem proves essentially that the subspace topology of a subspace topology is a subspace topology.

Theorem 12.8. *Let $\langle X, \mathcal{T} \rangle$ be a topological space and $Z \subseteq Y \subseteq X$ then $(\mathcal{T}_Y)_Z = \mathcal{T}_Z$*

Proof. First we have

$$
\begin{aligned}
U \in (\mathcal{T}_Y)_Z &\Rightarrow \exists V \in \mathcal{T}_Y \text{ such that } U = V \bigcap Z \\
&\Rightarrow \exists W \in \mathcal{T} \text{ such that } V = W \bigcap Y \\
&\Rightarrow U = \left(W \bigcap Z\right) \bigcap Y = W \bigcap \left(Z \bigcap Y\right) \\
&\underset{Z \subseteq Y}{\Rightarrow} U = W \bigcap Z \\
&\Rightarrow U \in \mathcal{T}_Z
\end{aligned}
$$

Also we have

$$\begin{array}{rcl} U\in\mathcal{T}_Z & \Rightarrow & \exists W\in\mathcal{T}\text{ such that } U=W\bigcap Z\\ & \underset{Z\subseteq Y}{\Rightarrow} & U=W\bigcap(Y\bigcap Z)=(W\bigcap Y)\bigcap Z\\ & \underset{W\bigcap Y\in\mathcal{T}_Y}{\Rightarrow} & U\in(\mathcal{T}_Y)_Z \end{array}$$

$\square$

Definition 12.9. *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $A\in X$ then $A^\circ=\bigcup_{U\in\{V\in\mathcal{T}|U\subseteq A\}}U$ is called the inner set of A (note that $\emptyset\in\{V\in\mathcal{T}|V\subseteq A\}$). Note that A° is open (see (4) in the definition of a topology) and that $A^\circ\subseteq A$.*

Theorem 12.10. *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $A\in X$ then if U is a open set such that $U\subseteq A$ then $U\subseteq A^\circ$. So A° is the biggest open set contained in A.*

Proof. If U is open and $U\subseteq A$ then $U\in\{V\in\mathcal{T}|V\subseteq A\}$ so that $U\subseteq\bigcup_{V\in\{V\in\mathcal{T}|V\subseteq A\}}V=A^\circ$ $\square$

12.1.1 Closed sets

Definition 12.11. *Let $\langle X,\mathcal{T}\rangle$ be a topological space then $\mathcal{C}\subseteq\mathcal{P}(X)$ the set of closed sets is defined by $\mathcal{C}=\{A\in\mathcal{P}(X)|X\setminus A\in\mathcal{T}\}$*

Note 12.12. If $\langle X,\mathcal{T}\rangle$ is a topological space and A is a closed set then $X\setminus A$ is open

Proof. If A is open then $\exists U$ open such that $A=X\setminus U=X\bigcap U^c$ and then $X\setminus A=X\bigcap A^c=X\bigcap(X\bigcap U^c)^c=X\bigcap(X^c\bigcup U)=X\bigcap U=U$ open $\square$

The next theorem proves that we can also define a topology using closed sets

Theorem 12.13. *Let $\langle X,\mathcal{T}\rangle$ be a topological space then the set of closed sets $\mathcal{C}$ satisfies the following:*

1. $\emptyset\in\mathcal{C}$
2. $X\in\mathcal{C}$
3. *If $A,B\in\mathcal{C}$ then $A\bigcup B\in\mathcal{C}$*
4. *If $\{A_i\}_{i\in I}$ is a family in $\mathcal{C}$ (a family of closed sets) then $\bigcap_{i\in I}A_i\in\mathcal{C}$*

Furthermore when $\mathcal{C}\subseteq\mathcal{P}(X)$ is a set of subsets of X satisfying 1,2,3 and 4 of the above then $\mathcal{T}=\{X\setminus A|A\in\mathcal{C}\}$ then $\langle X,\mathcal{T}\rangle$ is a topology with the set of closed sets $\mathcal{C}$.

Proof.

1. $X\setminus\emptyset=X\in\mathcal{T}\Rightarrow\emptyset\in\mathcal{C}$

2. $X\setminus X=\emptyset\in\mathcal{T}\Rightarrow X\in\mathcal{C}$

3. If $A,B\in\mathcal{C}$ then $X\setminus(A\bigcup B)\underset{1.31}{=}(X\setminus A)\bigcap(X\setminus B)\in\mathcal{T}$ (as $X\setminus A, X\setminus B\in\mathcal{T}$ and the definition of a topology).

4. If $\{A_i\}_{i\in I}$ is a family in $\mathcal{C}$ then $X\setminus(\bigcap_{i\in I}A_i)\underset{1.108}{=}\bigcup_{i\in I}(A\setminus A_i)\in\mathcal{T}$ (as $\forall i\in I$ we have $X\setminus A_i\in\mathcal{T}$)

Assume now that $\mathcal{C}\subseteq\mathcal{P}(X)$ fulfills 1,2,3 and 4 and define $\mathcal{T}=\{U\in\mathcal{P}(X)|X\setminus U\in\mathcal{C}\}$ then

1. $\emptyset=X\setminus X\Rightarrow\emptyset\in\mathcal{T}$

2. $X=X\setminus\emptyset\Rightarrow X\in\mathcal{T}$

3. If $U,V\in\mathcal{T}$ then $\exists A,B\in\mathcal{C}$ such that $U=X\setminus A$, $V=X\setminus B$ and $U\bigcap V=(X\setminus A)\bigcap(X\setminus B)=X\setminus(A\bigcup B)\in\mathcal{T}$ (as $A\bigcup B\in\mathcal{C}$)

4. If $\{U_i\}_{i\in I}$ is a family in $\mathcal{T}$ then $\forall i\in I$ there exists $A_i\in\mathcal{C}$ such that $U_i=X\setminus A_i$ and thus $\bigcup_{i\in I}U_i=\bigcup_{i\in I}(X\setminus A_i)\underset{1.108}{=}X\setminus(\bigcap_{i\in I}A_i)$

Now if A is closed in $\mathcal{T}$ then $X\setminus A\in\mathcal{T}$ so there exists a $C\in\mathcal{C}$ such that $X\setminus A=X\setminus C$ we have then $A\underset{A\subseteq X\wedge 1.31}{=}X\setminus(X\setminus A)\underset{X\setminus A=X\setminus C}{=}X\setminus(X\setminus C)\underset{C\subseteq X\wedge 1.31}{=}C$ and thus the set of closed sets of $\mathcal{T}$ is a subset of $\mathcal{C}$. If $A\in\mathcal{C}$ then by definition $X\setminus A\in\mathcal{T}$ so A is closed in $\boldsymbol{\mathcal{T}}$ and thus $\mathcal{C}$ is a subset of the set of closed sets in $\mathcal{T}$. These two last remarks proves that $\mathcal{C}$ is indeed the set set of closed sets of $\mathcal{T}$. □

Theorem 12.14. *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $A\subseteq X$ then the set of closed sets of $\langle A,\mathcal{T}_A\rangle$ is $\{C\bigcap A|C\text{ is closed in }\langle X,\mathcal{T}\rangle\}$*

Proof. If B is closed in $\langle A,\mathcal{T}_A\rangle$ then $A\setminus B\in\mathcal{T}_A\Rightarrow\exists U\in\mathcal{T}$ such that $(A\setminus B)=U\bigcap A$ then $B\underset{B\subseteq A\wedge 1.31}{=}B=A\setminus(A\setminus B)=A\setminus(U\bigcap A)=(A\setminus U)\bigcup(A\setminus A)=(A\setminus U)\bigcup\emptyset=A\setminus U=A\bigcap U^c\underset{A\subseteq X}{=}(A\bigcap X)\bigcap U^c=A\bigcap(X\bigcap U^c)=A\bigcap(X\setminus U)=(X\setminus U)\bigcap A\in\{C\bigcap A|C\text{ is closed in }\mathcal{T}\}$. If $B\in\{C\bigcap A|C\text{ is closed in }\langle X,\mathcal{T}\rangle\}$ then there exists a $C\subseteq X$ such that $X\setminus C\in\mathcal{T}$ and $B=C\bigcap A$. Now $A\setminus B=A\setminus(C\bigcap A)=(A\setminus C)\bigcup(A\setminus A)=A\setminus C\underset{A\in X}{=}(A\bigcap X)\setminus C=(A\bigcap X)\bigcap C^c=A\bigcap(X\bigcap C^c)=A\bigcap(X\setminus C)\in\mathcal{T}_A$ as $X\setminus C\in\mathcal{T}$ so we have that B is closed in $\langle A,\mathcal{T}_A\rangle$ □

Definition 12.15. *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $A\subseteq X$ then $\bar{A}=\bigcap_{C\in\{C|C\text{ is closed and }A\subseteq C\}}C$ is closed (as a intersection of a family of closed sets), contains A and is called the closure of A (note that $X\in\{C|C\text{ is closed and }A\subseteq C\}\Rightarrow\{C|C\text{ is closed and }A\subseteq C\}\neq\emptyset$)*

Theorem 12.16. *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $A\subseteq X$ then $\bar{A}$ is the smallest closed set containing A*

Proof. If C is a closed set such that $A\subseteq C$ then $C\in\{C|C\text{ is closed and }A\subseteq C\}$ and thus $\bigcap_{B\in\{C|C\text{ is closed and }A\subseteq C\}}B\subseteq C$ □

Theorem 12.17. *If $\langle X, \mathcal{T}\rangle$ is a topological space then $A \subseteq X$ is closed if and only of $A = \bar{A}$*

Proof.

($\Rightarrow$) If A is closed then as $A \subseteq A$ we have that $A \subseteq \bar{A} \subseteq A \Rightarrow A = \bar{A}$

($\Leftarrow$) If $A = \bar{A}$ then as $\bar{A}$ is closed we have that A is closed □

Definition 12.18. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and $A \subseteq X$ then $x \in X$ is a limit point or accumulation point of A iff $\forall U \in \mathcal{T}$ with $x \in U$ we have $\emptyset \neq (A \setminus \{x\}) \bigcap U$. The set of all limit points of A is called the derived set of A and is denoted by A'.*

Theorem 12.19. *Let $\langle X, \mathcal{T}\rangle$ is a topological space and $A \subseteq X$ then $\bar{A} = A \bigcup A'$*

Proof. If $x \in \bar{A}$ then we have two cases

1. **($x \in A$)** then trivially $x \in A \bigcup A'$
2. **($x \notin A$)** then if $x \notin A'$ then there exists a $U \in \mathcal{T}$ with $x \in U$ such that $\emptyset = (A \setminus \{x\}) \bigcap U \underset{x \in A \Rightarrow A = A \setminus \{x\}}{=} A \bigcap U \Rightarrow \forall a \in A \subseteq X$ we have $a \notin U \wedge a \in X \Rightarrow A \subseteq X \setminus U$ which means as $X \setminus U$ is closed that $\bar{A} \subseteq X \setminus U \Rightarrow x \notin U$ contradicting $x \in U$. So we must have $x \in A' \Rightarrow x \in A \bigcup A'$

So we have proved that $\bar{A} \subseteq A \bigcup A'$.

If now $x \in A \bigcup A'$ then we have either

1. **($x \in A$)** then as $A \subseteq \bar{A}$ we have $x \in \bar{A}$
2. **($x \notin A$)** then we must have $x \in A'$. If $x \notin \bar{A}$ then $x \in X \setminus \bar{A}$ which is open and thus we have $\emptyset \neq (A \setminus \{x\}) \bigcap (X \setminus \bar{A}) \underset{x \notin A \Rightarrow A \setminus \{x\} = A}{=} A \bigcap (X \setminus \bar{A})$ but then there exists a $y \in A \wedge y \notin \bar{A} \underset{y \in A \subseteq \bar{A}}{\Rightarrow} y \in \bar{A} \wedge y \notin \bar{A}$ a contradiction. So we must have $x \in \bar{A}$

This proves then $A \bigcup A' \subseteq \bar{A}$ and thus finally $\bar{A} = A \bigcup A'$ □

Corollary 12.20. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and $A \subseteq X$ then $x \in \bar{A}$ if only if $\forall U \in \mathcal{T}$* with *$x \in U$ we have $U \bigcap A \neq \emptyset$. Or in other words $\bar{A} = \{x \in X | \forall U \in \mathcal{T} \vdash x \in U$* we have *$A \bigcap U \neq \emptyset\}$*

Proof. If $x \in \bar{A} = A \bigcup A'$ then either

1. **($x \in A$)** then trivially $\forall U \in \mathcal{T}$ with $x \in U$ we have $x \in U \bigcap A \Rightarrow U \bigcap A \neq \emptyset$
2. **($x \in A'$)** then $\forall U \in \mathcal{T}$ with $x \in U$ we have $\emptyset \neq (A \setminus \{x\}) \bigcap U \Rightarrow U \bigcap A \neq \emptyset$

so in all cases we have $\forall U \in \mathcal{T}$ with $x \in U$ we have $U \bigcap A \neq \emptyset$.

On the other hand if $x \in X$ is such that $\forall U \in \mathcal{T}$ with $x \in U$ we have that $U \bigcap A \neq \emptyset$. If now $x \notin \bar{A}$ then $x \in X \setminus \bar{A}$ a open set so that we have $\emptyset \neq (X \setminus \bar{A}) \bigcap A = (X \bigcap \bar{A}^c) \bigcap A = X \bigcap (\bar{A}^c \bigcap A) \underset{A \subseteq \bar{A}}{=} X \bigcap \emptyset = \emptyset \Rightarrow \emptyset \neq \emptyset$ a contradiction. So we must have that $x \in \bar{A}$ □

As A is closed if and only if $A = \bar{A}$ (see 12.17) we have the following

Corollary 12.21. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and $A \subseteq X$ then A is closed if and only if $A = \{x \in X | \forall U \in \mathcal{T} \vdash x \in U \text{ we have } A \bigcap U \neq \emptyset\}$*

Definition 12.22. *Let X be a set and $\mathcal{T}_1$, $\mathcal{T}_2$ two topologies for X then $\mathcal{T}_2$ is finer then $\mathcal{T}_1$ iff $\mathcal{T}_1 \subseteq \mathcal{T}_2$. So if $\mathcal{T}_1$ is finer then $\mathcal{T}_2$ and $\mathcal{T}_1$ is finer then $\mathcal{T}_2$ then $\mathcal{T}_1 = \mathcal{T}_2$.*

12.1.2 Basis of a topological space

Definition 12.23. *Let $\langle X, \mathcal{T}\rangle$ be a topological space then $\mathcal{B} \subseteq \mathcal{T}$ is a basis for the topology $\mathcal{T}$ if every open set is the union of a family in $\mathcal{B}$. In other words $\mathcal{B}$ is a basis for $\mathcal{T}$ iff $\mathcal{T} = \{U \in \mathcal{P}(X) | \exists \{B_i\}_{i \in I} \text{ in } \mathcal{B} \text{ such that } U = \bigcup_{i \in I} U_i\}$*

Example 12.24. If $U = \emptyset$ then $\{B_i\}_{i \in \emptyset}$ (a empty family) is such that $U = \bigcup_{i \in \emptyset} B_i$ [if $x \in \bigcup_{i \in \emptyset} B_i$ then $\exists i \in \emptyset$ such that $x \in B_i$ which is impossible for the empty set, so $\emptyset = \bigcup_{i \in \emptyset} B_i$].

Theorem 12.25. *Let $\langle X, \mathcal{T}\rangle$ be a topological space with a basis $\mathcal{B}$ then if $A \subseteq X$ we have that $\mathcal{B}_A = \{B \bigcap A | B \in \mathcal{B}\}$ is a basis for the subspace topology $\mathcal{T}_A$ of A*

Proof. Let $U \in \mathcal{T}_A$ then there exists a $V \in \mathcal{T}$ such that $U = V \bigcap A$. By definition of a basis there exists a family $\{B_i\}_{i \in I}$ in $\mathcal{B}$ (meaning $\forall i \in I$ we have $B_i \in \mathcal{B}$) such that $V = \bigcup_{i \in I} B_i$ and thus $U = V \bigcap A = (\bigcup_{i \in I} B_i) \bigcap A = \bigcup_{i \in I} (B_i \bigcap A)$ so $V = \bigcup_{i \in I} (B_i \bigcap A)$ where $\{B_i \bigcap A\}_{i \in I}$ is a family in $\mathcal{B}_A$. Finally as $\mathcal{B} \subseteq \mathcal{T}$ we have $\forall B \in \mathcal{B}$ that $B \bigcap A \in \mathcal{T}_A$ and thus that $\mathcal{B}_A \subseteq \mathcal{T}_A$. □

Theorem 12.26. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and let $\mathcal{B} \subseteq \mathcal{T}$ then $\mathcal{B}$ is a basis for $\mathcal{T}$ if and only if $\forall U \in \mathcal{T}$ and $\forall x \in U$ we have that $\exists B \in \mathcal{B}$ such that $x \in B \subseteq U$*

Proof.

($\Rightarrow$) If $U \in \mathcal{T}$ then there exists a $\{B_i\}_{i \in I}$ in $\mathcal{B}$ such that $U = \bigcup_{i \in I} B_i$ so that if $x \in U$ then there exists a $i \in I$ with $x \in B_i \subseteq \bigcup_{i \in I} B_i \subseteq U$ where of course $B_i \in \mathcal{B}$

($\Leftarrow$) If $U \in \mathcal{T}$ then $\forall x \in U$ there exists a $B_x \in \mathcal{B}$ such that $x \in B_x \subseteq U$ this forms a family $\{B_x\}_{x \in U}$ in $\mathcal{B}$ such that $\bigcup_{x \in U} B_x = U$ proving that $\mathcal{B}$ is a basis for $\mathcal{T}$ □

Theorem 12.27. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and let $\mathcal{B}$ be a basis for $\mathcal{T}$ then*

1. *$\forall x \in X$ we have $\exists B \in \mathcal{B}$ with $x \in B$*
2. *$\forall B_1, B_2 \in \mathcal{B}$ then if $x \in B_1 \bigcap B_2$ then $\exists B_3 \in \mathcal{B}$ such that $x \in B_3 \subseteq B_1 \bigcap B_2$*

Proof.

1. As $X \in \mathcal{T}$ there exists a $\{B_i\}_{i \in I}$ in $\mathcal{B}$ such that $X = \bigcup_{i \in I} B_i$ and thus if $x \in X$ there exists a $i \in I$ such that $x \in B_i \in \mathcal{B}$

2. If $x \in B_1 \bigcap B_2$ then as $B_1, B_2 \in \mathcal{B} \subseteq \mathcal{T} \Rightarrow B_1 \bigcap B_2$ there exists a $\{C_i\}_{i \in I}$ in $\mathcal{B}$ such that $B_1 \bigcap B_2 = \bigcup_{i \in I} C_i$ so if $x \in B_1 \bigcap B_2$ then there exists a $i \in I$ such that $x \in C_i \subseteq \bigcup_{i \in I} C_i = B_1 \bigcap B_2$ and thus $x \in C_i \subseteq B_1 \bigcap B_2$. □

The above show the necessary condition that a basis must full fill, the following shows that any set of subset full filling the above conditions can be the basis of a topology.

Theorem 12.28. *Let X be as set and let $\mathcal{B} \subseteq \mathcal{P}(X)$ be a set of subsets of X full filling*

1. *$\forall x \in X$ there exists a $B \in \mathcal{B}$ such that $x \in B$*
2. *$\forall B_1, B_2 \in \mathcal{B}$ then if $x \in B_1 \bigcap B_2$ there exists a $B_3 \in \mathcal{B}$ such that $x \in B_3 \subseteq B_1 \bigcap B_2$*

then

$$\mathcal{T} = \{U \in \mathcal{P}(X) | \forall x \in U \text{ we have } \exists B \in \mathcal{B} \text{ such that } x \in B \subseteq U\}$$

is a topology for X that has $\mathcal{B}$ as its basis. $\mathcal{B}$ is called the generating basis for $\mathcal{T}$ and $\mathcal{T}$ is called the topology generated by $\mathcal{B}$

Proof. First we prove that $\mathcal{T}$ is a topology for X

a) $\emptyset \in \mathcal{T}$ (because every element in the empty set full fills vacuously every condition.

b) $X \in \mathcal{T}$ for if $x \in X$ then by (1) there exists a $B \in \mathcal{B}$ such that $x \in B \subseteq X$

c) If $\{U_i\}_{i \in I}$ is a family in $\mathcal{T}$ then if $x \in \bigcup_{i \in I} U_i$ there exists a $i \in I$ such that $x \in U_i \in \mathcal{T}$ so there exists a $B \in \mathcal{B}$ such that $x \in B \subseteq U_i \subseteq \bigcup_{i \in I} U_i$ and thus $\bigcup_{i \in I} U_i \in \mathcal{T}$

d) If $U_1, U_2 \in \mathcal{T}$ then if $x \in U_1 \bigcap U_2 \Rightarrow x \in U_1 \wedge x \in U_2$ so there exists $B_1, B_2 \in \mathcal{B}$ such that $x \in B_1 \subseteq U_1 \wedge x \in B_2 \subseteq U_2 \Rightarrow x \in B_1 \bigcap B_2 \subseteq U_1 \bigcap U_2$ and by (2) there exists a B_3 such that $x \in B_3 \subseteq B_1 \bigcap B_2 \subseteq U_1 \bigcap U_2$ proving that $U_1 \bigcap U_2 \in \mathcal{T}$

So $\langle X, \mathcal{T} \rangle$ is a topological space.

Also if $B \in \mathcal{B}$ then if $\forall x \in B$ there exists a $B \in \mathcal{B}$ such that $x \in B \subseteq B$ and thus $B \in \mathcal{T}$ giving $\mathcal{B} \subseteq \mathcal{T}$.

Next if $U \in \mathcal{T}$ and if $x \in U$ then by definition there exists a $B \in \mathcal{B}$ such that $x \in B \subseteq U$ which by 12.26 means that $\mathcal{B}$ is a basis for $\mathcal{T}$. □

The following theorem shows how to create a topology out of a set of topologies

Theorem 12.29. *Let X be a set and $\{\langle X_i, \mathcal{T}_i \rangle\}_{i \in I}$ a family of topological spaces such that $\forall i \in I$ we have $X_i \subseteq X$, $X = \bigcup_{i \in I} X_i$ and $\forall i, j \in I$, $U \in \mathcal{T}_i$, $V \in \mathcal{T}_j$ if $x \in U \bigcap V$ then $\exists W \in \mathcal{T}_i \bigcap \mathcal{T}_j$ such that $x \in W \subseteq U \bigcap V$. Then $\mathcal{B} = \bigcup_{i \in I} \mathcal{T}_i$ full fills the requirement for a generating basis (see 12.28) so that $\langle X, \mathcal{T} \rangle$ with $\mathcal{T} = \{U \subseteq X | \forall x \in U \text{ we have } \exists B \in \mathcal{B} \text{ such that } x \in B \subseteq U\}$ the generated topology. Also $\forall i \in I$ we have that $\mathcal{T}_i = \mathcal{T}_{X_i}$ (the subspace topology for X_i). We call this procedure to create a topology from subspace topologies **stitching topologies**.*

Proof. First we prove that the requirements of 12.28 are satisfied:

1. If $x\in X=\bigcup_{i\in I}X_i\Rightarrow\exists i\in I\vdash x\in X_i \underset{\mathcal{T}_i \text{ is a topology}\Rightarrow X_i\in\mathcal{T}_i}{\Rightarrow} \exists U\in\mathcal{T}_i\subseteq\mathcal{B}$ such that $x\in U$.
2. If $B_1, B_2\in\mathcal{B}$ and $x\in B_1\bigcap B_2$ then $\exists i,j\in I$ with $B_1\in\mathcal{T}_i$, $B_2\in\mathcal{T}_j$ then $\exists W\in\mathcal{T}_i\bigcap\mathcal{T}_j\subseteq\mathcal{B}$ such that $x\in W\subseteq B_1\bigcap B_2$

By 12.28 we have then that $\langle X,\mathcal{T}\rangle$ is a topological vector space.

Now if $i\in I$ and $U\in\mathcal{T}_{X_i}$ then there exists a $V\in\mathcal{T}$ such that $U=V\bigcap X_i$. Then $\forall x\in U$ we have $x\in V$ and thus $\exists B_x\in\mathcal{B}$ such that $x\in B_x\subseteq V$, but then there exists a $i_x\in I$ such that $B_x\in\mathcal{T}_{i_x}$. We have then that $V=\bigcup_{x\in V}B_x$ where $B_x\in\mathcal{T}_{i_x}$ and thus $U=(\bigcup_{x\in V}B_x)\bigcap X_i=\bigcup_{x\in V}(B_x\bigcap X_i)$. If now $y\in B_x\bigcap X_i$ where $B_x\in\mathcal{T}_{i_x}$ and $X_i\in\mathcal{T}_i$ then by the hypothesis there exists a $W_y\in\mathcal{T}_{i_x}\bigcap\mathcal{T}_i\subseteq\mathcal{T}_i$ such that $y\in W_y\subseteq B_x\bigcap X_i$ so that $B_x\bigcap X_i=\bigcup_{y\in B_x\bigcap X_i}W_y\in\mathcal{T}_i$ [as $\mathcal{T}_i$ is a topology and $W_y\in\mathcal{T}_i$]. So as $B_x\bigcap X_i\in\mathcal{T}_i$ we have that $U=V\bigcap X_i=\bigcup_{x\in V}(B_x\bigcap X_i)\in\mathcal{T}_i\Rightarrow U\in\mathcal{T}_i$ giving

$$\mathcal{T}_{X_i}\subseteq\mathcal{T}_i \tag{12.1}$$

If now $U\in\mathcal{T}_i\subseteq\mathcal{B}\subseteq\mathcal{T}$ then as $U\subseteq X_i$ we have $U=U\bigcap X_i\in\mathcal{T}_{X_i}$ so that we have

$$\mathcal{T}_i\subseteq\mathcal{T}_{X_i} \tag{12.2}$$

Using 12.1 and 12.2 we have finally $\mathcal{T}_i=\mathcal{T}_{X_i}$ □

Theorem 12.30. *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $A\subseteq X$ then*

1. *$x\in\bar{A}\Leftrightarrow\forall U\in\mathcal{T}\vdash x\in U$ we have $U\bigcap A\neq\emptyset$*
2. *If $\mathcal{B}$ is a basis of the topology $\mathcal{T}$ then $x\in\bar{A}\Leftrightarrow\forall U\in\mathcal{B}$ with $x\in U$ we have $U\bigcap A\neq\emptyset$*

Proof.

1. *This is already proved in 12.20*
2. *If $x\in\bar{A}$ then if $U\in\mathcal{B}\subseteq\mathcal{T}$ is such that $x\in U$ we have by (1) that $U\bigcap A\neq\emptyset$. If $x\in X$ is such that $\forall B\in\mathcal{B}\vdash x\in B$ we have $B\bigcap A\neq\emptyset$ then if $U\in\mathcal{T}$ is such that $x\in U$ there exists a $B\in\mathcal{B}$ such that $x\in B\subseteq U$ and then $\emptyset\neq B\bigcap A\subseteq U\bigcap A\Rightarrow U\bigcap A\neq\emptyset$ so that by (1) we have $x\in\bar{A}$.* □

Theorem 12.31. *Let $\langle X,\mathcal{T}\rangle$ be a topological space then if $A\subseteq X$ is closed we have*

1. *$x\in A\Leftrightarrow\forall U\in\mathcal{T}\vdash x\in U$ we have $U\bigcap A\neq\emptyset$*
2. *If $\mathcal{B}$ is a basis of the topology $\mathcal{T}$ then $x\in A\Leftrightarrow\forall U\in\mathcal{B}$ with $x\in U$ we have $U\bigcap A\neq\emptyset$*

Proof. This is trivial using 12.17 and the previous theorem 12.19. □

Theorem 12.32. *Let X be a set and $\mathcal{T}_1$, $\mathcal{T}_2$ two topologies on X with basis $\mathcal{B}_1$, $\mathcal{B}_2$ then the following is equivalent*

1. *$\mathcal{T}_2$ is finer then $\mathcal{T}_1$*

2. $\forall x \in X$, $\forall B \in \mathcal{B}_1 \vdash x \in B$ *there* $\exists B' \in \mathcal{B}_2$ *such that* $x \in B' \subseteq B$

Proof.

1. **(1 ⇒ 2)** then if $x \in X$ and $B \in \mathcal{B}_1 \subseteq \mathcal{T}_1 \subseteq \mathcal{T}_2$ such that $x \in B$ then there exists a $B' \in \mathcal{B}_2$ such that $x \in B' \subseteq B$.
2. **(2 ⇒ 1)** Let $U \in \mathcal{T}_1$ then $\forall x \in U$ there exists a $B_x \in \mathcal{B}_1$ such that $x \in B_x \subseteq U$ and thus by (2) $\exists B'_x \in \mathcal{B}_2$ such that $x \in B'_x \subseteq B_x \subseteq U$ and thus $U = \bigcup_{x \in U} B'_x \in \mathcal{T}_2$ proving that $\mathcal{T}_1 \subseteq \mathcal{T}_2$ □

Theorem 12.33. *Let X be a set and $\mathcal{S} \subseteq \mathcal{P}(X)$ then the set $\mathcal{B} = \{B \in \mathcal{P}(S) | \exists \{S_i\}_{i \in I}$ a finite non empty family in $\mathcal{S}$ such that $B = \bigcap_{i \in I} S_i\} \bigcup \{X\}$ satisfies the conditions to become a basis (see 12.28) and generates thus a topology $\mathcal{T} = \{U \in \mathcal{P}(X) | \forall x \in U \text{ there} \exists B \in \mathcal{B} \text{such that} \, x \in B \subseteq U\}$. This topology is called the topology generated by the sub-base $\mathcal{S}$. $\mathcal{B}$ is called the basis generated by the sub basis.*

Proof. We have to show that $\mathcal{B}$ satisfies the conditions from 12.28

1. Let $x \in X$ then $x \in X \in \mathcal{B}$
2. Let $x \in B_1 \bigcap B_2$ where $B_1, B_2 \in \mathcal{B}$ then there exists finite families $\{S_i\}_{i \in I_1}$ and $\{T_i\}_{i \in I_2}$ such that $B_1 = \bigcap_{i \in I_1} S_i$ and $B_2 = \bigcap_{i \in I_2} T_i$. Form then the family $\{R_i\}_{i \in (I_1 \times \{0\}) \bigcup (I_2 \times \{1\})}$ by $R_{(i,0)} = S_i$ and $R_{(i,1)} = T_i$ then we have

$$
\begin{aligned}
x \in \bigcap_{i \in (I_1 \times \{0\}) \bigcup (I_2 \times \{1\})} R_i &\Leftrightarrow \forall (i,j) \in (I_1 \times \{0\}) \bigcup (I_2 \times \{1\}) \text{ we have } x \in R_{(i,j)} \\
&\Leftrightarrow [\forall (i,j) \in I_1 \times \{0\} \text{ we have } x \in R_{(i,j)}] \wedge [\forall (i,j) \in I_2 \times \{1\} \text{ we have } x \in R_{(i,j)}] \\
&\Leftrightarrow [\forall i \in I_1 \text{ we have } x \in R_{(i,0)}] \wedge [\forall i \in I_2 \text{ we have } x \in R_{(i,1)}] \\
&\Leftrightarrow [\forall i \in I_1 \text{ we have } x \in S_i] \wedge [\forall i \in I_2 \text{ we have } x \in T_i] \\
&\Leftrightarrow x \in \bigcap_{i \in I_1} S_i \wedge x \in \bigcap_{i \in I_2} T_i \\
&\Leftrightarrow x \in B_1 \bigcap B_2
\end{aligned}
$$

This proves that $B_1 \bigcap B_2 = \bigcap_{i \in (I_1 \times \{0\}) \bigcup (I_2 \times \{1\})} R_i \in \mathcal{B}$ and thus we have found a $B = B_1 \bigcap B_2 \in \mathcal{B}$ such that $x \in B \subseteq B_1 \bigcap B_2$. □

Definition 12.34. *Let I be a non empty set, $\{\langle X_i, \mathcal{T}_i \rangle\}_{i \in I}$ be a family of topological spaces then the box topology $\mathcal{T}_{\text{box}}$ on $\prod_{i \in I} X_i$ is the topology generated by the basis $\mathcal{B} = \{\prod_{i \in I} U_i | \forall i \in I \text{ we have } U_i \in \mathcal{T}_i\}$.*

Proof. We have of course prove that 12.28 is satisfied.

1. If $x \in \prod_{i \in I} X_i$ then as $\forall i \in I$ we have $X_i \in \mathcal{T}_i \Rightarrow \prod_{i \in I} X_i \in \mathcal{B}$
2. Let $U_1 = \prod_{i \in I} V_i \in \mathcal{B}$, $U_2 = \prod_{i \in} W_i \in \mathcal{B}$ then $\forall i \in I$ we have $V_i, W_i \in \mathcal{T}_i \Rightarrow V_i \bigcap W_i \in \mathcal{T}_i$ we have then $U_1 \bigcap U_2 \underset{2.95}{=} \prod_{i \in I} (V_i \bigcap W_i) \in \mathcal{B}$ [as $V_i \bigcap W_i \in \mathcal{T}_i$] □

Theorem 12.35. *Let I be a non empty set, $\{\langle X_i, \mathcal{T}_i\rangle\}_{i\in I}$ be a family of topological spaces so that $\forall i\in I$ we have that $\mathcal{B}_i$ is a basis of $\mathcal{T}_i$. Then $\mathcal{B}=\{\prod_{i\in I} B_i|\forall i\in I\vDash B_i\in \mathcal{B}_i\}$ is a basis for the box topology on $\prod_{i\in I} X_i$.*

Proof. Let $U\in\mathcal{T}_{\text{box}}$ and $x\in U$ then as $\{\prod_{i\in I} U_i|\forall i\in I\vDash U_i\in\mathcal{T}_i\}$ we have by 12.26 that there exists a $\{U_i\}_{i\in I}$ where $\forall i\in I\vDash U_i\in\mathcal{T}_i$ such that $x\in\prod_{i\in I} U_i\subseteq U$. Now $\forall i\in I$ we have $x_i\in U_i$ and thus as $\mathcal{B}_i$ is a basis for $\mathcal{T}_i$ there exists a $B\in\mathcal{B}_i$ such that $x_i\in B\subseteq U_i$. So if we define $\forall i\in I$ $\mathcal{A}_i=\{B\in\mathcal{B}_i|x_i\in B\subseteq U_i\}$ then $\mathcal{A}_i\neq\emptyset$ and thus by the axiom of choice (see 2.201) there exists a choice function $B: I\to\bigcup_{i\in I}\mathcal{A}_i$ such that $\forall i\in I$ we have $B_i\in\mathcal{A}_i$ and thus $x_i\in B_i\subseteq U_i$ and $B_i\in\mathcal{B}_i$. This defines a family $\{B_i\}_{i\in I}$ such that $\forall i\in I$ we have $B_i\in\mathcal{B}_i$ and $x_i\in B_i\subseteq U_i$ and thus by 2.80 and 2.85 we have $x\in\prod_{i\in I} B_i\subseteq\prod_{i\in I} U_i\subseteq U$ where $\prod_{i\in I} B_i\in\mathcal{B}$, using 12.26 again we have then proved that $\mathcal{B}$ is a basis for the box topology. □

Definition 12.36. *Let I be a non empty set, $\{\langle X_i,\mathcal{T}_i\rangle\}_{i\in I}$ be a family of topological spaces then the product topology $\mathcal{T}_{\text{product}}$ on $\prod_{i\in I} X_i$ is defined by the topology generated by the sub base $\mathcal{S}=\{\pi_i^{-1}(V)|i\in I\wedge V\in\mathcal{T}_i\}$*

Theorem 12.37. *Let I be a non empty set, $\{\langle X_i,\mathcal{T}_i\rangle\}_{i\in I}$ be a family of topological spaces then the product topology $\mathcal{T}_{\text{product}}$ on $X=\prod_{i\in I} X_i$ has as its basis $\mathcal{B}=\{\prod_{i\in I} U_i|$*where $\{U_i\}_{i\in I}$ is *a* family such that $\exists A$ finite $\subseteq I$ such that $\forall i\in I\setminus A$ we have $U_i=X_i$ and $\forall i\in A$ we have $U_i\in\mathcal{T}_i\}$

Proof. First $\mathcal{S}=\{\pi_i^{-1}(V)|i\in I,\ V\in\mathcal{T}_i\}$ is the generating sub basis of $\mathcal{T}_{\text{product}}$ and the basis $\mathcal{B}'$ for $\mathcal{T}_{\text{product}}$ is given by $\mathcal{B}'=\{B\in\mathcal{P}(S)|\exists\{S_i\}_{i\in J}$ a finite non empty family in $\mathcal{S}$ such that $B=\bigcap_{i\in J} S_i\}\bigcup\{X\}$. We have then:

1. $(\boldsymbol{B\in\mathcal{B}'})$ then we have either

   a. $(\boldsymbol{B=X})$ then $B=\prod_{i\in I} X_i$ and if we take $A=\emptyset$ (which is finite) then $\forall i\in A$ we have $X_i\in\mathcal{T}_i$ vacuously and $\forall i\in I=I\setminus A$ we have $X_i\in\mathcal{T}_i$ and thus $B\in\mathcal{B}'$

   b. $(\boldsymbol{\exists\{S_i\}_{i\in J}}$ **a finite non empty family in** $\boldsymbol{\mathcal{S}}$ **such that** $\boldsymbol{B=\bigcap_{i\in J} S_i})$ as $\forall i\in J$ we have $S_i\in\mathcal{S}$ there exists a $k_i\in I$ and $U_i\in\mathcal{T}_{k_i}$ such that $S_i=\pi_{k_i}^{-1}(U_i)$. Now if we take $\{B_{k_i,k}\}_{k\in I}$ where $B_{k_i,k}=\left\{\begin{smallmatrix} U_i \text{ if } k=k_i\\ X_k \text{ if } k\neq k_i\end{smallmatrix}\right.\Rightarrow$ $B_{k_i,k}\in\mathcal{T}_k$[if $k=k_i$ then $B_{k_i,k}=U_i\in\mathcal{T}_{k_i}=\mathcal{T}_k$ and if $k\neq k_i$ then $B_{k_i,k}=X_k\in\mathcal{T}_k$]. Then we have $S_i=\pi_{k_i}^{-1}(U_i)=\prod_{k\in I} B_{k_i,k}$ [If $x\in\pi_{k_i}^{-1}(U_i)$ then $\pi_{k_i}(x)\in U_i=B_{k_i,k_i}$ and $\forall k\in I\setminus\{k_i\}$ we have $\pi_k(x)\in X_k$ so that $x\in\prod_{k\in I} B_{k_i,k}$, if $x\in\prod_{k\in I} B_{k_i,k}$ then $\pi_{k_i}(x)\in B_{k_i,k_i}=U_i\Rightarrow x\in\pi_{k_i}^{-1}(U_i)$] So $B=\bigcap_{i\in J}(\prod_{k\in I} B_{k_i,k})\underset{2.96}{=}\prod_{k\in I}(\bigcap_{i\in J} B_{k_i,k})=\prod_{k\in I} C_k$ where $\forall k\in I$ we have that $C_k=\bigcap_{i\in J} B_{k_i,k}$. As J is finite and $B_{k_i,k}\in\mathcal{T}_k$ we have that $C_k\in\mathcal{T}_k$. Define now $A\subseteq I$ by $A=\{k_i|i\in J\}$ which is finite as J is finite. If $k\in A$ we have $C_k\in\mathcal{T}_k$ and if $k\in I\setminus A$ then $\forall i\in J$ we have $k_i\neq k\Rightarrow B_{k_i,k}=X_k\Rightarrow C_k=X_k$. This proves that $B\in\mathcal{B}$.

2. **($B\in\mathcal{B}$)** then there exists a family $\{U_i\}_{i\in I}$ and a finite $A\subseteq I$ such that $\forall i\in A$ we have $U_i\in\mathcal{T}_i$ and $\forall i\in I\setminus A$ we have $U_i=X_i$ with $B=\prod_{i\in I}U_i$. Define now $\{U_{i,j}\}_{i\in A\times I}$ by $A_{i,j}=\begin{cases}U_i \text{ if } i=j\\ X_j \text{ if } i\neq j\end{cases}$ then $U_i=\bigcap_{j\in A}A_{j,i}$ [if $x\in U_i\subseteq X_i$ then $\forall j\in A$ we have either $i=j\Rightarrow x\in U_i=A_{i,i}$ or if $i\neq j$ then $x\in X_i=A_{j,i}$ so that $x\in\bigcap_{j\in A}A_{j,i}$, if $x\in\bigcap_{j\in A}A_{j,i}$ then if $i\in I\setminus A$ we have $\forall j\in A$ that $j\neq i$ so that $A_{j,i}=X_i=U_i$ and $x\in U_i$ and if $i\in A$ we have $x\in A_{i,i}=U_i$]. So $B=\prod_{i\in I}\left(\bigcap_{j\in A}A_{j,i}\right)\underset{2.96}{=}\bigcap_{j\in A}\left(\prod_{i\in I}A_{j,i}\right)$. Now $\forall j\in A$ we have $\prod_{i\in I}A_{j,i}=\pi_j^{-1}(U_i)$ [if $x\in\prod_{i\in I}A_{j,i}$ then $\pi_j(x)\in A_{j,j}=U_j\Rightarrow x\in\pi_j^{-1}(U_j)$, if $x\in\pi_j^{-1}(U_j)$ then $x\in X$ such that $\pi_j(x)\in U_j=A_{j,j}$ and if $i\neq j$ we have $\pi_i(x)\in X_i=A_{j,i}$ so that $x\in\prod_{i\in I}A_{j,i}$] so that $B=\bigcap_{j\in A}\pi_j^{-1}(U_j)$ proving that $B\in\mathcal{B}'$ □

Corollary 12.38. *Let I be a non empty set, $\{\langle X_i,\mathcal{T}_i\rangle\}_{i\in I}$ be a family of topological spaces then the product topology $\mathcal{T}_{\text{product}}$ on $X=\prod_{i\in I}X_i$ has as its basis $\mathcal{B}=\{\bigcap_{i\in A}\pi_i^{-1}(U_i)|\mathcal{A}\subseteq I\wedge\mathcal{A}\text{ is finite and }\forall i\in\mathcal{A}\text{ we have that }U_i\text{ is open in }X_i\}$*

Proof. Take $\mathcal{B}$ the basis of X defined in the previous theorem (see 12.37) then $B\in\mathcal{B}$ if and only if $\exists\mathcal{A}\subseteq I$, $\mathcal{A}$ finite and $B=\prod_{i\in I}U_i$ where $U_i=X_i$ if $i\in I\setminus\mathcal{A}$ and U_i is open in X_i if $i\in\mathcal{A}$. We prove now that $B=\bigcap_{i\in\mathcal{A}}\pi_i^{-1}(U_i)$

1. If $x\in\prod_{i\in I}U_i$ then if $i\in\mathcal{A}$ we haven that $\pi_i(x)\in U_i$ or $x\in\pi_i^{-1}(U_i)\Rightarrow x\in\bigcap_{i\in\mathcal{A}}\pi_i^{-1}(U_i)$
2. If $x\in\bigcap_{i\in\mathcal{A}}\pi_i^{-1}(U_i)$ then $\forall i\in\mathcal{A}$ we have $x\in\pi_i^{-1}(U_i)\Rightarrow x_i=\pi_i(x)\in U_i$. If $i\in I\setminus\mathcal{A}$ then we have trivially that $\pi_i(x)\in X_i=U_i$ so that $x\in\prod_{i\in I}U_i$ □

Using the above theorem it is now easy to prove that in the finite case the box topology is equivalent with the product topology.

Theorem 12.39. *Let $\{\langle X_i,\mathcal{T}_i\rangle_{i\in I}\}$ be a family of topological spaces then $\mathcal{T}_{\text{box}}$ on $X=\prod_{i\in I}X_i$ is finer then $\mathcal{T}_{\text{product}}$ on X. If I is finite then $\mathcal{T}_{\text{box}}=\mathcal{T}_{\text{product}}$*

Proof. Let $\mathcal{B}_{\text{product}}\underset{12.37}{=}\{\prod_{i\in I}U_i|\text{where }\{U_i\}_{i\in I}\text{ is }a\text{ family such that }\exists A\text{ finite}\subseteq I\text{ such that }\forall i\in I\setminus A\text{ we have }U_i=X_i\text{ and }\forall i\in A\text{ we have }U_i\in\mathcal{T}_i\}$ be the basis for the product topology and let $\mathcal{B}_{\text{box}}=\{\prod_{i\in I}U_i|\forall i\in I\text{ we have }U_i\in\mathcal{T}_i\}$ be the basis of the box topology. Then if $B\in\mathcal{B}_{\text{product}}$ there exists a family $\{U_i\}_{i\in I}$ and a finite $A\subseteq I$ such that $\forall i\in A$ we have $U_i\in\mathcal{T}_i$, $\forall i\in I\setminus A$ we have $U_i=X_i\in\mathcal{T}_i$ and $B=\prod_{i\in I}U_i$ so that $B\in\mathcal{B}_{\text{box}}$, this proves that $\mathcal{B}_{\text{product}}\subseteq\mathcal{B}_{\text{box}}$ and thus by 12.32 we have that $\mathcal{T}_{\text{box}}$ is finer then $\mathcal{T}_{\text{product}}$ or $\mathcal{T}_{\text{product}}\subseteq\mathcal{T}_{\text{box}}$. If I is finite then if $B\in\mathcal{B}_{\text{box}}$ then there exists a family $\{U_i\}_{i\in I}$ such that $\forall i\in I$ finite we have $U_i\in\mathcal{T}_i$ and $\forall i\in I\setminus I=\emptyset$ we have $U_i=X_i$ vacuously with $B=U_i$ proving that $B\in\mathcal{B}_{\text{product}}$. This proves that $\mathcal{B}_{\text{product}}=\mathcal{B}_{\text{box}}$ and thus that $\mathcal{T}_{\text{product}}=\mathcal{T}_{\text{box}}$. □

Theorem 12.40. *Let $\{\langle X_i,\mathcal{T}_i\rangle\}_{i\in I}$ be a family of topological spaces and let $\forall i\in I$ $A_i\subseteq X_i$ be equipped with the subspace topology $\mathcal{T}_{A_i}$ then the subspace topology of $\prod_{i\in I}A_i\subseteq\prod_{i\in I}X_i$ is the same as the product topology of $\{\langle A_i,\mathcal{T}_{A_i}\rangle\}_{i\in I}$*

Proof. Let $\mathcal{B}_1$ be the base of the product topology on $\prod_{i\in I} X_i$ then the subspace topology on $\prod_{i\in I} A_i$ is defined by $\mathcal{B}_1 = \{B\bigcap(\prod_{i\in I} A_i)|B\in\mathcal{B}_1\}$ and let $\mathcal{B}_2$ be a basis for the product topology of $\prod_{i\in I} A_i$.

1. If $U\in\mathcal{B}_1 \underset{12.38}{\Rightarrow} U = (\prod_{i\in I} U_i)\bigcap(\prod_{j\in I} A_j)$ where $U_i = X_i$ except for a finite subset of I where $U_i\in\mathcal{T}_i \Rightarrow U = \prod_{i\in I}(U_i\bigcap A_i)$ where $(U_i\bigcap A_i) = A_i$ (if $U_i = X_i$) except for a finite subset of I where $(U_i\bigcap A_i)\in\mathcal{T}_{A_i}$ and thus $U\in\mathcal{B}_2$
2. If $U\in\mathcal{B}_2$ then $B = \prod_{i\in I} U_i$ where $U_i = A_i = A_i\bigcap A_i$ except for a finite subset of I where $U_i = W_i\bigcap A_i$ ($W_i\in\mathcal{T}_i$) and thus if $V_i = X_i$ except for a finite subset where $V_i = W_i$ then $B = \prod_{i\in I}(V_i\bigcap A_i) = (\prod_{i\in I} V_i)\bigcap(\prod_{i\in I} A_i)$ where $V_i = A_i$ except for a finite subset of I where $V_i\in\mathcal{T}_i \Rightarrow B\in\mathcal{B}_1$

now as we just have proved that $\mathcal{B}_1 = \mathcal{B}_2$ we use 12.32 to prove that both topologies are the same. □

12.1.3 Dense sets

Definition 12.41. *Let $\langle X, \mathcal{T}\rangle$ be a toplogical space then $A\subseteq X$ is a dense subset if $\bar{A} = X$*

Definition 12.42. *A topological space $\langle X, \mathcal{T}\rangle$ is a Baire space if for every family $\{A_i\}_{i\in\mathbb{N}}$ with $\forall i\in\mathbb{N}$ we have that A_i is closed and $A_i^\circ = \emptyset$, we have that $(\bigcup_{i\in\mathbb{N}} A_i)^\circ = \emptyset$. In other words the union of any family of closed sets with empty interior has also empty interior.*

Theorem 12.43. *Let $\langle X, \mathcal{T}\rangle$ be a topological space, $A\subseteq X$ then $A^\circ = \emptyset$ if and only if $X\setminus A$ is dense in A*

Proof.

1. **($\Rightarrow$)** If $A^\circ = \emptyset$ then if $x\in X$ and U open with $x\in U$ we have $U\nsubseteq A$ (otherwise $x\in A^\circ$)$\Rightarrow U\bigcap(X\setminus A)\neq\emptyset\Rightarrow\overline{X\setminus A} = X$ proving that $X\setminus A$ is dense in X.
2. **($\Leftarrow$)** Assume $X\setminus A$ is dense in A and that $x\in A^\circ$ which as A° is open means by the density of $X\setminus A$ in X means that $A^\circ\bigcap(X\setminus A)\neq\emptyset$ so there exists a $y\notin A$ and $y\in A^\circ\subseteq A$ which is a contradiction. This means that $A^\circ = \emptyset$. □

Theorem 12.44. *Let $\langle X, \mathcal{T}\rangle$ be a topological space then $\langle X, \mathcal{T}\rangle$ is a Baire space if and only if for every $\{U_i\}_{i\in\mathbb{N}}$ with $\forall i\in\mathbb{N}$ we have that U_i is open and dense in X then $\bigcap_{i\in\mathbb{N}} U_i$ is dense in X. In other words the intersection of every sequence of open dense sets in X is dense in X.*

Proof.

1. **($\Rightarrow$)** Let $\{U_i\}_{i\in\mathbb{N}}$ be a family of open dense sets in X define then $\{A_i\}_{i\in\mathbb{N}}$ where $\forall i\in\mathbb{N}$ we have $A_i = X\setminus U_i$ a closed set. As $U_i = X\setminus(X\setminus U_i) = X\setminus A_i$ is by assumption dense in X we have by the previous theorem that $\forall i\in\mathbb{N}$ we have $A_i^\circ = \emptyset$. By definition of a Baire space w have then that $(\bigcup_{i\in\mathbb{N}} A_i)^\circ = \emptyset$ and thus by the previous theorem we have $X\setminus(\bigcup_{i\in\mathbb{N}} A_i)$ is dense in X. Also $X\setminus(\bigcup_{i\in\mathbb{N}} A_i) = X\bigcap(\bigcup_{i\in\mathbb{N}} A_i)^c = X\bigcap(\bigcap_{i\in\mathbb{N}} A_i^c) = \bigcap_{i\in\mathbb{N}} X\bigcap A_i^c = \bigcap_{i\in\mathbb{N}}(X\setminus A_i) = \bigcap_{i\in\mathbb{N}} U_i$ so that $\bigcap_{i\in\mathbb{N}} U_i$ is dense in X.

2. **($\Leftarrow$)** Let $\{A_i\}_{i\in\mathbb{N}}$ a family of closed sets with $A_i^\circ = \emptyset$ for all $i \in \mathbb{N}$, define then $\{U_i\}_{i\in\mathbb{N}}$ by $\forall i \in \mathbb{N}\ U_i = X \setminus A_i$. As A_i is closed we have that U_i is open and by the previous theorem we have that U_i is dense in X. By the assumption we have then that $\bigcap_{i\in\mathbb{N}} U_i$ is dense in X. Using the previous theorem we have then that $\emptyset = (X \setminus \bigcap_{i\in\mathbb{N}} U_i)^\circ = (X \bigcap (\bigcap_{i\in\mathbb{N}} U_i)^c)^\circ = (X \bigcap (\bigcup_{i\in\mathbb{N}} U_i^c))^\circ = (\bigcup_{i\in\mathbb{N}} (X \bigcap U_i^c))^\circ = (\bigcup_{i\in\mathbb{N}} (X \bigcap (X \bigcap A_i^c)^c))^\circ = (\bigcup_{i\in\mathbb{N}} (X \bigcap (X^c \bigcup A_i)))^\circ = (\bigcup_{i\in\mathbb{N}} (X \bigcap A_i))^\circ \underset{A_i \subseteq X}{=} (\bigcup_{i\in\mathbb{N}} A_i)^\circ$ proving that $\langle X, \mathcal{T}\rangle$ is Baire. □

12.2 Metric Spaces

Definition 12.45. *A pseudo metric space $\langle X, d\rangle$ is a non empty set together with a function $d: X \times X \rightarrow \mathbb{R}$ [d is called the pseudo metric] such that $\forall x, y, z \in X$ we have*

1. $d(x, x) = 0$
2. $d(x, y) = d(y, x)$
3. $d(x, z) \leqslant d(x, y) + d(y, z)$

If also $d(x, y) = 0 \Rightarrow x = y$ then we call d a metric and $\langle X, d\rangle$ a metric space

Theorem 12.46. *If $\langle X, d\rangle$ is a pseudo metric space then $\forall x, y \in X$ we have $d(x, y) \geqslant 0$ (distance are positive).*

Proof. If $x, y \in X$ then $0 = d(x, x) \leqslant d(x, y) + d(y, x) = d(x, y) + d(x, y) = 2 \cdot d(x, y) \Rightarrow 0 \leqslant d(x, y)$ □

Definition 12.47. *Let $\langle X, d_X\rangle$ and $\langle Y, d_Y\rangle$ be pseudo metric spaces and $f: X \rightarrow Y$ a **partial function** (should not be a function but can be a function if $\mathrm{dom}(f) = X$). Given a $x \in X$ and a $y \in Y$ we say that $\lim_{h\rightarrow x} f(h) = y$ if and only if $\forall \varepsilon > 0\ (\varepsilon \in \mathbb{R})$ there exists a $\delta > 0$ ($\delta \in \mathbb{R}$) such that $\forall x' \in \mathrm{dom}(f)$ with $0 < d_X(x, x') < \varepsilon$ we have that $d_Y(y, f(x')) < \varepsilon$*

Note 12.48. Note that neither x should be in $\mathrm{dom}(f)$ or y should be in $\mathrm{range}(f)$

Definition 12.49. *Given a pseudo metric space $\langle X, d\rangle$ and $\varepsilon \in \mathbb{R}_+$ then $B_d(x, \varepsilon) = \{y \in X | d(x, y) < \varepsilon\}$ is called the open ball centered around x with radius ε.*

Definition 12.50. *Given a pseudo metric space $\langle X, d\rangle$ and $\varepsilon \in \mathbb{R}_+$ then $\bar{B}_d(x, \varepsilon) = \{y \in X | d(x, y) \leqslant \varepsilon\}$ is called the closed ball centered around x with radius ε.*

Remark 12.51. Until further notice we adopt the usual convention that $\varepsilon > 0$ means the same as $\varepsilon \in \mathbb{R}_+$

Theorem 12.52. *Given a pseudo metric space $\langle X, d\rangle$ then if $x \in B_d(x_1, \varepsilon_1) \bigcap B_d(x_2, \varepsilon_2)$ then $\exists \varepsilon > 0$ such that $x \in B_d(x, \varepsilon) \subseteq B_d(x_1, \varepsilon_1) \bigcap B_d(x_2, \varepsilon_2)$*

Proof. Let $x \in B_d(x_1, \varepsilon_1) \bigcap B_d(x_2, \varepsilon_2) \Rightarrow d(x_1, x) < \varepsilon_1 \wedge d(x_2, x) < \varepsilon$, take then $\varepsilon = \min(\varepsilon_1 - d(x_1, x), \varepsilon_2 - d(x_2, x)) > 0$ then if $y \in B_d(x, \varepsilon)$ we have $d(x, y) < \varepsilon$ so for $i = 1, 2$ we have $d(x_i, y) \leqslant d(x_i, x) + d(x, y) < d(x_i, x) + \varepsilon \leqslant d(x_i, x) + (\varepsilon_i - d(x_i, x)) = \varepsilon_i \Rightarrow y \in B_d(x_i, \varepsilon_i)$ □

As $y \in B_d(x, \varepsilon)$ means $y \in B_d(x, \varepsilon) \bigcap B_d(x, \varepsilon)$ we can use the previous theorem to prove the following corollary.

Corollary 12.53. *Given a pseudo metric space $\langle X, d\rangle$ and $y \in B_d(x, \varepsilon)$ then $\exists \delta > 0$ such that $B_d(y, \delta) \subseteq B_d(x, \varepsilon)$*

Theorem 12.54. *Given a pseudo metric space $\langle X, d\rangle$ we have that $\mathcal{B} = \{B_d(x, \varepsilon) | \varepsilon \in \mathbb{R}_+ \wedge x \in X\}$ satisfies the requirements for a generating basis (see 12.28), the generated topology on X is called the topology generated by the pseudo metric and is noted by $\mathcal{T}_d$.*

Proof. To see if $\mathcal{B}$ satisfies the requirements of 12.28 note that:

1. $\forall x \in X$ we have $x \in B_d(x, 1) \in \mathcal{B}$
2. $\forall B_1, B_2 \in \mathcal{B}$ with $x \in B_1 \bigcap B_2$ we have $B_1 = B_d(x_1, \varepsilon_1)$, $B_2 = B_d(x_2, \varepsilon_2)$ then as $x \in B_d(x_1, \varepsilon_1) \bigcap B_d(x_2, \varepsilon_2)$ we have by 12.52 the existence of a δ such that $x \in B_d(x, \delta) \subseteq B_d(x_1, \varepsilon_1) \bigcap B_d(x_2, \varepsilon_2)$ so we found a $B_3 = B_d(x, \delta) \in \mathcal{B}$ such that $x \in B_3 \subseteq B_1 \bigcap B_2$ □

Corollary 12.55. *Given a pseudo metric space $\langle X, d\rangle$ then U is a open set in the metric topology if and only if $\forall x \in U$ there exists a $\delta > 0$ such that $x \in B_d(x, \delta)$*

Proof.

1. **($\Rightarrow$)** If $x \in U$ open in the metric topology then by the definition of a basis there exists a $y \in X$ and a $\varepsilon > 0$ such that $x \in B_d(y, \varepsilon) \subseteq U$. Using 12.53 there exists a $\delta > 0$ such that $x \in B_d(x, \delta) \subseteq B_d(y, \varepsilon) \subseteq U$
2. **($\Leftarrow$)** If for every $x \in U$ there exists a $\delta_x > 0$ such that $x \in B_d(x, \delta_x) \subseteq U \Rightarrow U = \bigcup_{x \in U} B_d(x, \delta_x)$ which is open as $B_d(x, \delta_x)$ is open by the definition of a basis.

□

Theorem 12.56. *Let $\langle X, d\rangle$ be a pseudo metric space then closed balls are indeed closed*

Proof. Let $\bar{B}_d(x, \varepsilon)$ be a closed ball and let $y \in X \setminus \bar{B}_d(x, \varepsilon) \Rightarrow d(x, y) > \varepsilon$ take then $\delta = d(x, y) - \varepsilon > 0$, now let $z \in B_d(y, \delta)$ then assume that $d(z, x) \leqslant \varepsilon \Rightarrow d(x, y) \leqslant d(x, z) + d(z, y) \leqslant \varepsilon + d(z, y) < \varepsilon + \delta = \varepsilon + d(x, y) - \varepsilon = d(x, y) \Rightarrow d(x, y) < d(x, y)$ a contradiction, so $d(z, x) > \varepsilon \Rightarrow z \in X \setminus \bar{B}_d(x, \varepsilon) \Rightarrow B_d(y, \delta) \subseteq X \setminus \bar{B}_d(x, \varepsilon) \underset{12.55}{\Rightarrow} X \setminus \bar{B}_d(x, \varepsilon)$ is open in the metric topology and thus $\bar{B}_d(x, \varepsilon)$ is closed □

Theorem 12.57. *Let $\langle X, d\rangle$ be a (pseudo) metric space and $A \subseteq X$ then the subspace topology on A is generated by the (pseudo) metric $d_{|A \times A}: A \times A \to \mathbb{R}$ [$d_{|A \times A}$ is the restriction of d to $A \times A$]*

Proof.

First we prove that $d_{|A\times A}$ is a pseudo metric so let $x, y, z \in A$ then we have

1. $d_{|A\times A}(x, x) = d(x, x) = 0$
2. $d_{|A\times A}(x, y) = d(x, y) = d(y, x) = d_{|A\times A}(y, x)$
3. $d_{|A}(x, y) = d(x, y) \leqslant d(x, z) + d(y, z) = d_{|A\times A}(x, z) + d_{|A\times A}(z, y)$
4. If d is a metric then we have if $d_{|A\times A}(x, y) = 0 \Rightarrow d(x, y) = 0 \underset{d \text{ is a metric}}{\Rightarrow} x = y$

Second $B_{d_{|A\times A}}(x, \delta) = \{y \in A | d_{|A\times A}(x, y) < \delta\} = \{y \in A | d(x, y) < \delta\} = \{y \in X | d(x, y) < \delta\} \bigcap A = B_d(x, \delta) \bigcap A$. Let now $\mathcal{T}$ the topology generated by d and $\mathcal{T}_A$ the subspace topology. Then if $V \in \mathcal{T}_A$ there exists a $U \in \mathcal{T}$ such that $V = U \bigcap A$, if now $x \in V$ then $x \in U$ and by 12.55 there exists a $\delta_x > 0$ such that $x \in B_d(x, \delta_x) \subseteq U$ and thus $x \in B_{d_{|A\times A}}(x, \delta_x) = B_d(x, \delta_x) \bigcap A \subseteq U \bigcap A = V$ so we have $x \in B_{d|A\times A}(x, \delta_x) \subseteq V$ and thus by 12.55 we conclude that V is open in the topology generated by $d_{|A\times A}$. Assume now that V is open in the topology generated by $d_{|A\times A}$ on A then $V \subseteq A$ and $\forall x \in V$ there exists a $\delta_x > 0$ such that $x \in B_{d_{|A\times A}}(x, \delta_x) \subseteq V$ and thus $x \in B_{d_{|A\times A}}(x,\delta_x) = B_d(x, \delta_x) \bigcap A \subseteq V$ so that $V = \bigcup_{x\in V} B_{d_{|A\times A}}(x, \delta_x)$ Take now the open set $U = \bigcup_{x\in V} B_d(x, \delta_x)$ in $\mathcal{T}$ then $U \bigcap A = (\bigcup_{x\in V} B_d(x, \delta_x)) \bigcap A = \bigcup_{x\in V} (B_d(x, \delta_x) \bigcap A) = V$ proving that V is open in the subspace topology. □

Definition 12.58. *Two pseudo metrics d_1, d_2 on a set X are equivalent iff they generate the same topology.*

Theorem 12.59. *Let d_1, d_2 be two metrics on a set X and let $\mathcal{T}_1$, $\mathcal{T}_2$ the generated metric topologies then $\mathcal{T}_2$ is finer then $\mathcal{T}_1$ if and only if $\forall x \in X, \forall \varepsilon > 0$ there $\exists \delta > 0$ such that $x \in B_{d_2}(x, \delta) \subseteq B_{d_1}(x, \varepsilon)$*

Proof.

1. **($\Rightarrow$)** If $x \in X$ and $\varepsilon > 0$ then by 12.32 for $x \in B_{d_1}(x, \varepsilon)$ there exists a $y \in X$, $\delta' > 0$ such that $x \in B_{d_2}(y, \delta') \subseteq B_{d_1}(x, \varepsilon)$, using 12.53 there exists a $\delta > 0$ such that $x \in B_{d_2}(x, \delta) \subseteq B_{d_2}(x, \delta') \subseteq B_{d_1}(x, \varepsilon)$
2. **($\Leftarrow$)** If $y \in B_{d_1}(x, \varepsilon)$ then using 12.53 there exists a $\delta' > 0$ such that $y \in B_{d_1}(y, \delta') \subseteq B_{d_1}(x, \varepsilon)$ and by the hypothesis there exists a $\delta > 0$ such that $y \in B_{d_2}(y, \delta) \subseteq B_{d_1}(y, \delta') \subseteq B_{d_1}(x, \varepsilon)$ proving by 12.32 that $\mathcal{T}_2$ is finer then $\mathcal{T}_1$ □

Definition 12.60. *Let $\langle X, d_X\rangle$ and $\langle Y, d_Y\rangle$ be two pseudo metric spaces then a function $\varphi: X \to Y$ is called a isometry iff*

1. *φ is a bijective map*
2. *$\forall x, y \in X$ we have $d_X(x, y) = d_Y(\varphi(x), \varphi(y))$*

Theorem 12.61. *Let $\langle X, d_X\rangle$, $\langle Y, d_Y\rangle$ be two pseudo metric spaces and $\varphi: X \to Y$ a isometry then φ^{-1} is a isometry.*

Proof. Because 2.38 we have that φ^{-1} is a bijection and if $x, y \in Y$ then $d_Y(x, y) = d_Y(\varphi(\varphi^{-1}(x)), \varphi(\varphi^{-1}(y))) \underset{d \text{ is isometry}}{=} d_X(\varphi^{-1}(x), \varphi^{-1}(y))$ □

Theorem 12.62. *Let $\langle X, d_X\rangle$, $\langle Y, d_Y\rangle$, $\langle Z, d_Z\rangle$ be metric spaces and let $\varphi: X \rightarrow Y$ and $\psi: Y \rightarrow Z$ be isometries then $\psi \circ \varphi: X \rightarrow Z$ is a isometry*

Proof. Using 2.46 we have that $\psi \circ \varphi'' X \rightarrow Z$ is a bijection and $d_X(x, y) = d_Y(\varphi(x), \varphi(y)) = d_Z(\psi(\varphi(x)), \psi(\varphi(y))) = d_Z((\psi \circ \varphi)(x), (\psi \circ \varphi)(y))$ □

Theorem 12.63. *Let $\langle X, d_X\rangle$, $\langle Y, d_Y\rangle$ be two metric spaces with a isometry $\varphi: X \rightarrow Y$ between them then $\varphi(B_{d_X}(x, \delta)) = B_{d_Y}(\varphi(x), \delta)$ and $\varphi^{-1}(B_{d_Y}(y, \delta)) = B_{d_X}(\varphi^{-1}(y), \delta)$*

Proof.

1. If $y \in \varphi(B_{d_X}(x, \delta)) \Rightarrow y = \varphi(z) \wedge z \in B_{d_X}(x, \delta) \Rightarrow y = \varphi(z) \wedge d_X(x, z) < \delta \Rightarrow y = \varphi(z) \wedge d_Y(\varphi(x), \varphi(z)) < \delta \Rightarrow d_Y(\varphi(x), y) < \delta \Rightarrow y \in B_{d_Y}(\varphi(x), \delta)$. If $y \in B_{d_Y}(\varphi(x), \delta) \Rightarrow d_Y(\varphi(x), y) < \delta \Rightarrow d_Y(\varphi(x), \varphi(\varphi^{-1}(y))) < \delta \Rightarrow d_X(x, \varphi^{-1}(y)) < \delta \Rightarrow \varphi^{-1}(y) \in B_{d_X}(x, \delta) \Rightarrow y \in \varphi(B_{d_X}(x, \delta))$. This proves that $\varphi(B_{d_X}(x, \delta)) = B_{d_Y}(\varphi(x), \delta)$.
2. Let $x \in \varphi^{-1}(B_{d_Y}(y, \delta)) \Rightarrow \varphi(x) \in B_{d_Y}(y, \delta) \Rightarrow d_Y(y, \varphi(x)) < \delta \Rightarrow d_Y(\varphi(\varphi^{-1}(y)), \varphi(x)) \Rightarrow d_X(\varphi^{-1}(y), x) < \delta \Rightarrow x \in B_{d_X}(\varphi^{-1}(y), \delta)$. If $x \in B_{d_X}(\varphi^{-1}(y), \delta) \Rightarrow d_X(\varphi^{-1}(y), x) < \delta \Rightarrow d_Y(\varphi(\varphi^{-1}(y)), \varphi(x)) < \delta \Rightarrow \varphi(x) \in B_{d_Y}(y, \delta) \Rightarrow x \in \varphi^{-1}(B_{d_Y}(y, \delta))$. This proves that $\varphi^{-1}(B_{d_Y}(y, \delta)) = B_{d_X}(\varphi^{-1}(y), \delta)$ □

Theorem 12.64. *Let $\langle X, d_X\rangle$, $\langle Y, d_Y\rangle$ be two metric spaces with a isometry $\varphi: X \rightarrow Y$ between then . If $\mathcal{T}_X$, $\mathcal{T}_Y$ be the two metric topologies on X, Y then $\mathcal{T}_X = \{\varphi^{-1}(V) | V \in \mathcal{T}_Y\} = \{U \subseteq X | \varphi(U) \in \mathcal{T}_Y\}$ and $\mathcal{T}_Y = \{\varphi(U) | U \in \mathcal{T}_X\} = \{V \subseteq Y | \varphi^{-1}(V) \in \mathcal{T}_X\}$.*

Proof. First we prove that

$$\{\varphi^{-1}(V) | V \in \mathcal{T}_Y\} = \{U \subseteq X | \varphi(U) \in \mathcal{T}_Y\}, \tag{12.3}$$

so if $U \in \{\varphi^{-1}(V) | V \in \mathcal{T}_Y)$ there exists a $V \in \mathcal{T}_Y$ such that $U = \varphi^{-1}(V) \subseteq X$ and as φ is a bijection we have $\varphi(U) = V \in \mathcal{T}_Y$ or $U \in \{U \subseteq X | \varphi(Y) \in \mathcal{T}_Y\}$ also if $U \in \{U \subseteq X | \varphi(U) \in \mathcal{T}_Y\}$ then $U \subseteq X \wedge \varphi(U) \in \mathcal{T}_Y$ and as φ is a bijection we have $U = \varphi^{-1}(\varphi(U))$ where $\varphi(U) \in \mathcal{T}_Y$ so that $U \in \{\varphi^{-1}(V) | V \in \mathcal{T}_Y\}$.

Second we prove that $\mathcal{T}_X = \{\varphi^{-1}(V) | V \in \mathcal{T}_Y\}$. So if $U \in \mathcal{T}_X$ take then $V = \varphi(U)$ and $y \in V$ so that $\exists x \in U$ such that $y = \varphi(x)$. Using 12.55 there exists a $\delta > 0$ such that $x \in B_{d_X}(x, \delta) \subseteq U$ and thus $y = \varphi(x) \in \varphi(B_{d_X}(x, \delta)) \subseteq \varphi(U) = V \underset{12.63}{\Rightarrow} y \in B_{d_Y}(\varphi(x), \delta) \subseteq V \underset{y = \varphi(x)}{\Rightarrow} y \in B_{d_Y}(y, \delta) \subseteq V$ which by 12.55 means that V is open in $\mathcal{T}_Y$ and as φ is bijective we have $U = \varphi^{-1}(V)$ where $V \in \mathcal{T}_Y$ such that $U \in \{\varphi^{-1}(V) | V \in \mathcal{T}_Y\}$ or

$$\mathcal{T}_X \subseteq \{\varphi^{-1}(V) | V \in \mathcal{T}_Y\} \tag{12.4}$$

Next if $U \in \{\varphi^{-1}(V) | V \in \mathcal{T}_Y\}$ then there exists a $V \in \mathcal{T}_Y$ such that $U = \varphi^{-1}(V)$ or as φ is a bijection we have $\varphi(U) = V$. If now $x \in U$ then $\varphi(x) \in \varphi(U) = V \in \mathcal{T}_Y$ and using 12.55 there exists a $\delta > 0$ such that $\varphi(x) \in B_{d_Y}(\varphi(x), \delta) \subseteq V = \varphi(U)$ so that $x \in \varphi^{-1}(B_{d_Y}(\varphi(x), \delta)) \subseteq \varphi^{-1}(\varphi(U)) \underset{\varphi\,\text{is a bijection}}{=} U \underset{12.63}{\Rightarrow} x \in B_{d_X}(\varphi^{-1}(\varphi(x)), \delta) \subseteq U \Rightarrow x \in B_{d_X}(x, \delta) \subseteq U$ proving by 12.55 that $U \in \mathcal{T}_X$ so that we have

$$\{\varphi^{-1}(V) | V \in \mathcal{T}_Y\} \subseteq \mathcal{T}_X \tag{12.5}$$

Using 12.4 and 12.5 we have then

$$\mathcal{T}_X = \{\varphi^{-1}(V) | V \in \mathcal{T}_Y\} \tag{12.6}$$

As for the rest note that $\Phi = \varphi^{-1}$ is a isometry from $Y \to X$ (see 12.61) so by 12.6 and 12.3 we have that $\mathcal{T}_Y = \{\Phi^{-1}(U) | U \in \mathcal{T}_X\} = \{V \subseteq Y | \Phi(V) \in \mathcal{T}_X\} \underset{\Phi = \varphi^{-1} \Rightarrow \Phi^{-1} = \varphi}{=} \{\varphi(U) | U \in \mathcal{T}_X\} = \{V \subseteq Y | \varphi^{-1}(V) \in \mathcal{T}_X\}$ □

Definition 12.65. *Let $\langle X, d\rangle$ be a pseudo metric space and $A \subseteq X$ a subset, then A is called* ***bounded*** *iff $\exists M \in \mathbb{R}$ such that $\forall x, y \in X$ we have $d(x, y) \leq M$. If A is bounded then $\sup(\{d(x, y) | x, y \in A\})$ exists ($\mathbb{R}, \leqslant$ is conditional complete (see 9.43)) and is called the diameter of A noted as $\operatorname{diam}(A)$.*

Example 12.66. Let $\langle X, d\rangle$ be a metric space then every closed and open ball is bounded

Proof. Let $\overline{B_d}(a, \delta)$ be a closed ball then if $M = 2 \cdot \delta$ we have if $x, y \in \overline{B_d}(a, \delta)$ that $d(x, y) \leqslant d(x, a) + d(a, y) \leqslant \delta + \delta = 2 \cdot \delta$. As $B_d(a, \delta) \subseteq \overline{B_d}(a, \delta)$ we have also that $\overline{B_d}(a, \delta)$ is bounded. □

Theorem 12.67. *Let $\{\langle X_i, d_i\rangle\}_{i \in I}$ be a finite family of (pseudo) metric spaces then d: $(\prod_{i \in I} X_i) \times (\prod_{i \in I} X_i) \to \mathbb{R}$ defined by $(x, y) \to d(x, y) = \max(\{d_i(\pi_i(x), \pi_i(y)) | i \in I\})$ (see 5.50) is a (pseudo) metric space and the topology generated is the product topology $\mathcal{T}_{\text{product}}$ and the box topology $\mathcal{T}_{\text{box}}$. The metric d is called the product metric.*

Proof. First we have to prove that d is indeed a (pseudo) metric

1. $d(x, x) = \max(\{d_i(\pi_i(x), \pi_i(x)) | i \in I\}) = 0$ as $\forall i \in I$ we have $d_i(\pi_i(x), \pi_i(x)) = 0$
2. $d(x, y) = \max(\{d_i(\pi_i(x), \pi_i(y)) | i \in I|\}) = \max(\{d_i(\pi_i(y), \pi_i(x)) | i \in I|\}) = d(y, x)$
3. $d(x, y) = \max(\{d_i(\pi_i(x), \pi_i(y)) | i \in I\}) \leqslant_{2.164} \max(\{d_i(\pi_i(x), \pi_i(z)) + d_i(\pi_i(z), \pi_i(y)) | i \in I\}) \leqslant \max(\{\max(\{d_i(\pi_i(x), \pi_i(z)) | i \in I|\}) + d_i(\pi_i(z), \pi_i(y)) | i \in I\}) \leqslant \max(\max(\{d_i(\pi_i(x), \pi_i(z)) | i \in I\}) + \max(\{d_i(\pi_i(z), \pi_i(y)) | i \in I|\})) = \max(\{d_i(\pi_i(x), \pi_i(z))\}) + \max(\{d_i(\pi_i(z), \pi_i(y)) | i \in I|\}) = d(x, z) + d(y, z)$
4. If $\{\langle X_i d_i\rangle\}_{i \in I}$ are metric spaces then if $d(x, y) = 0$ we have $\forall i \in I$ that $d_i(\pi_i(x), \pi_i(y)) = 0 \Rightarrow \forall i \in I$ we have $\pi_i(x) = \pi_i(y) \Rightarrow x = y$

Next as I is finite we have by 12.39 that $\mathcal{T}_{\text{product}} = \mathcal{T}_{\text{box}}$. The generating basis for the box topology is $\mathcal{B}_{\text{box}} = \{\prod_{i \in I} U_i | U_i \text{ open in the metric topology} \mathcal{T}_i \text{ of } \langle X_i, d\rangle\}$, the generating basis for the metric topology is $\mathcal{B}_d = \{B_d(x, \delta) | x \in X_i, \delta > 0\}$. Let $x \in X$ and $B \in \mathcal{B}_{\text{box}}$ with $x \in B$ then $B = \prod_{i \in I} U_i$ where $U_i \in \mathcal{T}_i$. So if $x \in B$ then $\forall i \in I$ we have $\pi_i(x) \in U_i$ so that $\exists \delta_i > 0$ such that $\pi_i(x) \in B_{d_i}(\pi_i(x), \delta_i) \subseteq U_i$, take now $\delta = \min(\{\delta_i | i \in I\}) > 0$ then if $y \in B_d(x, \delta)$ we have $d(x, y) < \delta \Rightarrow \forall i \in I$ we have $d_i(\pi_i(x), \pi_i(y)) < \delta \leqslant \delta_i \Rightarrow \forall i \in I$ we have $\pi_i(y) \in B_{d_i}(\pi_i(x), \delta_i) \subseteq U_i \Rightarrow y \in \prod_{i \in I} U_i = B$ so that we have proved that $x \in B_d(x, \delta) \subseteq B$ or

$$\forall x \in X, \forall B \in \mathcal{B}_{\text{box}} \text{ there } \exists B' \in \mathcal{B}_d \text{ such that } x \in B' \subseteq B \tag{12.7}$$

If now $x \in X$ and $B \in \mathcal{B}_d$ with $x \in B$ then $\exists \delta' > 0, y \in X$ such that $x \in B_d(y, \delta') = B$ which means that $\exists \delta > 0$ such that $x \in B_d(x, \delta) \subseteq B_d(y) = B$. Take now $\prod_{i \in I} B_{d_i}(\pi_i(x), \delta) \in \mathcal{B}_{\text{box}}$ then if $z \in \prod_{i \in I} B_{d_i}(\pi_i(x), \delta)$ then $\forall i \in I$ we have $\pi_i(z) \in B_{d_i}(\pi_i(x), \delta) \Rightarrow d_i(\pi_i(x), \pi_i(z)) < \delta \Rightarrow d(x, z) = \max(\{d_i(\pi_i(x), \pi_i(z)) | i \in I\}) < \delta \Rightarrow z \in B_d(x, \delta) \subseteq B$. So we have proved that $x \in \prod_{i \in I} B_{d_i}(\pi_i(x), \delta) \subseteq B$ or

$$\forall x \in X, \forall B \in \mathcal{B}_d \text{ there } \exists B' \in \mathcal{B}_{\text{box}} \text{ such that } x \in B' \subseteq B \tag{12.8}$$

Using then 12.32 we have then that the metric topology is equal to the box topology and thus also the product topology. □

12.3 Normed spaces

Definition 12.68. *A pseudo normed space $\langle X, \|\|\rangle$ is a (real or complex) vector space X together with a function $\|\|: X \to \mathbb{R}$ such that*

1. $\forall x \in X$ *we have* $\|x\| \geqslant 0$
2. $\forall x \in X, \forall \alpha \in \mathbb{R}(\mathbb{C})$ *we have* $\|\alpha \cdot x\| = |\alpha| \cdot \|x\|$
3. $\forall x, y \in X$ *we have* $\|x + y\| \leqslant \|x\| + \|y\|$

if also $\|x\| = 0 \Rightarrow x = 0$ then $\langle X, \|\|\rangle$ is called a normed space.

Note 12.69. Let $\langle X, \|\|\rangle$ be a normed space then if $x = 0$ we have $\|x\| = 0$

Proof. If $x = 0$ then $x = 0 \cdot x \Rightarrow \|x\| = \|0 \cdot x\| = |0| \cdot x = 0$ □

Theorem 12.70. *Let $\langle X, \|\|\rangle$ be a pseudo normed space then*

1. $\forall x, y$ *we have*
$$|\|x\| - \|y\|| \leqslant \|x + y\| \tag{12.9}$$
2. *If I is a finite set, $\{a_i\}_{i \in I}$ then*
$$\left\|\sum_{i \in I} a_i\right\| \leqslant \sum_{i \in I} \|a_i\| \tag{12.10}$$

Proof.

1. $\|x\| = \|x + y - y\| \leqslant \|x + y\| + \|-y\| = \|x + y\| + \|y\| \Rightarrow \|x\| - \|y\| \leqslant \|x + y\|$ and $\|y\| = \|x + y - x\| \leqslant \|x + y\| + \|-x\| = \|x + y\| + \|x\| \Rightarrow \|y\| - \|x\| \leqslant \|x + y\|$. From this it follows that $|\|x\| - \|y\|| \leqslant \|x + y\|$.
2. If $I = \emptyset$ then we have that $\sum_{i \in I} x_i = 0$ so that trivially $\|\sum_{i \in I} a_i\| \leqslant \sum_{i \in I} \|a_i\|$. If $I \neq \emptyset$ we first prove that $\|\sum_{i=1}^n a_i\| \leqslant \sum_{i=1}^n \|a_i\|$ by induction. So let $S = \{n \in \mathbb{N} | \|\text{If } \{a_i\}_{i \in \{1,\ldots,n\}} \subseteq X \text{ then } \sum_{i=1}^n a_i\| \leqslant \sum_{i=1}^n \|a_i\|\}$ then we have

   ***n* = 1.** then if $\{a_i\}_{i \in \{1,\ldots,1\}} \subseteq X$ then $\|\sum_{i=1}^1 a_i\| = \|a_1\| = \sum_{i=1}^1 \|a_i\|$ proving that $1 \in \mathcal{S}$

$\boldsymbol{n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}}$. if $\{a_i\}_{i\in\{1,\ldots,n+1\}} \subseteq X$ then $\|\sum_{i=1}^{n+1} a_i\| = \|a_{n+1} + \sum_{i=1}^{n} a_i\| \leqslant \|a_{n+1}\| + \|\sum_{i=1}^{n} a_i\| \leqslant_{n\in\mathcal{S}} \|a_{n+1}\| + \sum_{i=1}^{n} \|a_i\| = \sum_{i=1}^{n+1} \|a_i\|$ proving that $n+1 \in \mathcal{S}$

Finally if I is a finite non empoty set then there exists a $n \in \mathbb{N}$ and a bijection $b: \{1,\ldots,n\} \rightarrow I$ such that $\sum_{i\in I} a_i = \sum_{i=1}^{n} a_{b_i}$ so that $\|\sum_{i\in I} a_i\| = \|\sum_{i=1}^{n} a_{b_i}\| \leqslant \sum_{i=1}^{n} \|a_{b_i}\| = \sum_{i\in I} \|a_i\|$ □

Example 12.71. $\langle \mathbb{R}, | \,|| $ and $\langle \mathbb{C}, | \,||$ where $\|x\| = |x|$ forms. In $\langle \mathbb{R}, \|\rangle$ the basis of the normed topology is $\mathcal{B} = \{]a,b[| a,b \in \mathbb{R} \wedge a < b\}$

Proof. See 8.66 and 9.33. For the case of $\langle \mathbb{R}, \|\rangle$ its basis is $\mathcal{B}_{||} = \{B_{||}(x,\varepsilon) | x \in \mathbb{R} \wedge \varepsilon \in \mathbb{R}_+\}$ then if $B \in \mathcal{B}_{||}$ there exists a $x \in \mathbb{R} \wedge \varepsilon \in \mathbb{R}_+$ such that $B = \{y \in \mathbb{R} | |x-y| < \varepsilon\} =]x-\varepsilon, x+\varepsilon[\subseteq \mathcal{B}$. If $B \in \mathcal{B}$ then there exists a $a,b \in \mathbb{R}$ with $a<b$ such that $B =]a,b[$ as $a<b$ take $\varepsilon = \frac{b-a}{2}$ and $x = a+\varepsilon$ then we have $y \in]x-\varepsilon, x+\varepsilon[\Leftrightarrow x-\varepsilon < y < x+\varepsilon \Leftrightarrow a+\varepsilon-\varepsilon < y < a+\varepsilon+\varepsilon \Leftrightarrow a < y < a+2\varepsilon = a+b-a = b$ proving that $B \in \mathcal{B}_{||}$. So we have that $\mathcal{B} = \mathcal{B}_{||}$ □

Example 12.72. As a application of the above example, $\forall x \in \mathbb{R}$ we have that $[x,\infty[= \{y \in \mathbb{R} | y \geqslant x\}$ is closed.

Proof. $\mathbb{R} \backslash [x,\infty[=]-\infty.x[= \{y | y < x\}$ then if $y \in]-\infty, x[\Rightarrow y < x$ then by density of the reals there exists a $\delta \in \mathbb{R}$ such that $y < \delta < x$, take then $\varepsilon = \delta - y$, if now $z \in B_{||}(y,\varepsilon)$ we have $|y-z| < \varepsilon$. If now $x \leqslant z \Rightarrow y < x \leqslant z \Rightarrow z - y = |y-z| < \delta - y \Rightarrow z < \delta < x$ giving the contradiction $x < x$, so we must have that $z < x$ proving that $B_{||}(y,\varepsilon) \in \mathbb{R} \backslash [x,\infty]$ and thus that $R \backslash [x,\infty[$ is open and thus that $[x,\infty[$ is closed. □

Every (pseudo) normed space defines a (pseudo) metric space

Theorem 12.73. *Let $\langle X, \|\|\rangle$ be a (pseudo) normed space then $\langle X, d_{\|\|}\rangle$ where $d_{\|\|}(x,y) = \|x-y\|$ is a (pseudo) metric space. The generated topology is called the (pseudo) normed topology and its basis is $\mathcal{B} = \{B_{d_{\|\|}}(x,\delta) | x \in X, \delta > 0\}$ where $B_{d_{\|\|}}(x,\delta) = \{y | \|x-y\| < \delta\} \underset{\text{defined}}{=} B_{\|\|}(x,\delta)$. We note the generated topology by $\mathcal{T}_{\|\|}$.*

Proof.

1. $d_{\|\|}(x,y) = \|x-y\| \geqslant 0$
2. $d_{\|\|}(x,y) = \|x-y\| = \|(-1)\cdot(y-x)\| = |-1| \cdot \|y-x\| = \|y-x\| = d(y,x)$
3. $d_{\|\|}(x,z) = \|x-z\| = \|x-y+y-z\| \leqslant \|x-y\| + \|y-z\| = d_{\|\|}(x,y) + d_{\|\|}(y, z)$
4. If $\|\|$ is normed then if $d(x,y) = 0 \Rightarrow \|x-y\| = 0 \Rightarrow x-y = 0 \Rightarrow x = y$ □

Example 12.74. If $\langle \mathbb{R}, \|\rangle$ is the normed vector spaces of real numbers then the normed topology is generated by $\mathcal{B} = \{]a,b[, a,b \in \mathbb{R}, a < b]$ where $]a,b[= \{x | a < x \wedge x < b\}$

Proof. The generating basis for the normed topology is $\mathcal{B}=\{B_{||}(x,\delta)|x\in\mathbb{R},\delta>0\}$ now $B\in\mathcal{B}\Rightarrow\exists x\in\mathbb{R}$ and $\delta>0$ such that $B=B_{||}(x,\delta)$ so if $y\in B$ we have $|x-y|<\delta$ then $x-y,y-x<|x-y|<\delta\Rightarrow x-y<\delta\wedge y-x<\delta\Rightarrow x-\delta<y\wedge y<x+\delta\Rightarrow y\in]x-\delta,x+\delta[$ proving that $B_{||}(x,\delta)\subseteq]x-\delta,x+\delta[$. If $y\in]x-\delta,x+\delta[\Rightarrow x-\delta<y\wedge y<x+\delta\Rightarrow x-y<\delta\wedge-(x-y)<\delta\Rightarrow|x-y|<\delta$ proving that $B_{||}(x,\delta)=]x-\delta,x+\delta[$. Finally if $a,b\in\mathbb{R}$ with $a<b$ then $0<\delta=\frac{b-a}{2}$ we have then $](a+\delta)-\delta,(a+\delta)+\delta[=]a,a+2\cdot\delta[=]a,a+b-a[=]a,b[$ □

Proposition 12.75. *Let $\langle X,\|\|\rangle$ be a (psuedo) normed space with generated topology $\mathcal{T}_{\|\|}$ and $A\subseteq X$ then if we define $\|\|_{|A}:A\rightarrow\mathbb{R}$ by $\|x\|_A=\|\|$ then $\langle A,\|\|_A\rangle$ is a (pseudo) normed space and $\mathcal{T}_{\|\|_A}$ is the subspace topology on A induced by $\mathcal{T}_{\|\|}$.*

Proof. First we prove that $\|\|_A$ is a pseudo norm on A

1. $\forall x\in A$ we have $\|x\|_A=\|x\|\geqslant 0$
2. $\forall x\in A$ and $\forall\alpha\in\mathbb{R}(\mathbb{C})$ we have $\|\alpha\cdot x\|_A=\|\alpha\cdot x\|=\alpha\cdot\|x\|=\alpha\cdot\|x\|_A$
3. $\forall x.y\in A$ we have $\|x+y\|_A=\|x+y\|\leqslant\|x\|+\|y\|=\|x\|_A+\|y\|_A$

Further if $\|\|$ is a norm then if $\|x\|_A=0$ we have $\|x\|=0$ proving that $x=0$ proving that in this case $\|\|_A$ is also a norm. Next as $\forall x,y\in A$ we have that $(d_{\|\|})_{|A\times A}(x,y)=d_{\|\|}(x,y)=\|x-y\|=\|x-y\|_A=d_{\|\|_A}(x,y)$ it follows that $(d_{\|\|})_{|A\times A}=d_{\|\|_A}$. Applying then 12.57 proves that $\mathcal{T}_{\|\|_A}$ s the subspace topology on A induced by $\mathcal{T}_{\|\|}$. □

Theorem 12.76. *Let $\langle X,\|\|\rangle$ be a normed vector space over $\mathbb{K}=(\mathbb{C}\text{ or }\mathbb{R})$ then $\forall\alpha\neq 0,\alpha\in\mathbb{K}$ and $\forall x\in X$ we have (see 10.120 for the definitions)*

1. *$\forall M\subseteq X$ we have $\overline{\alpha\cdot M}=\alpha\cdot\bar{M}$*
2. *$\forall M$ open in X we have that $\alpha\cdot M$ is open*
3. *$\forall M$ open in X we have that $x+M$ is open*
4. *$\forall M$ open ins X and $A\subseteq X$ we have that $A+M$ is open*
5. *$\forall M\subseteq X$ we have $\bar{M}-\bar{M}\subseteq\overline{M-M}$*

Proof.

1. If $x\in\overline{\alpha\cdot M}$ then as $\alpha\neq 0$ we can define $z=\frac{1}{\alpha}\cdot x$ so that $x=\alpha\cdot z$. If now $z\in U$ open then there exists a $\delta>0$ such that $z\in B_{\|\|}(z,\delta)\subseteq U$. Now as $x\in B_{\|\|}(x,|\alpha|\cdot\delta)$ a open set and $x\in\overline{\alpha\cdot M}$ there exists by 12.20 a $m\in M$ such that $\alpha\cdot m\in B_{\|\|}(x,|\alpha|\cdot\delta)$ or $\|x-\alpha\cdot m\|<|\alpha|\cdot\delta\Rightarrow\|\alpha\cdot z-\alpha\cdot m\|<|\alpha|\cdot\delta\Rightarrow|\alpha|\cdot\|z-m\|<|\alpha|\cdot\delta\Rightarrow\|z-m\|<\delta\Rightarrow m\in B_{\|\|}(z,\delta)\subseteq U$ so we have $U\bigcap M\neq\emptyset$. So using 12.20 we have that $z\in\bar{M}$ or that $x\in\alpha\cdot\bar{M}$, this proves that

$$\overline{\alpha\cdot M}\subseteq\alpha\cdot\bar{M}\tag{12.11}$$

If $x \in \alpha \cdot \bar{M}$ and $x \in U$ a open set, then there exists a $\delta > 0$ such that $x \in B_{\|\|}(x, \delta) \subseteq U$, now as $x \in \alpha \cdot \bar{M}$ there exists a $z \in \bar{M}$ such that $x = \alpha \cdot z$ and as $z \in B_{\|\|}\left(z, \frac{\delta}{|\alpha|}\right)$ a open set we have by 12.20 that there exist a $m \in M$ such that $m \in B_{\|\|}\left(z, \frac{\delta}{|\alpha|}\right)$ or $\|z - m\| < \frac{\delta}{|\alpha|} \Rightarrow |\alpha| \cdot \|z - m\| < \delta \Rightarrow \|\alpha \cdot z - \alpha \cdot m\| < \delta \Rightarrow \|x - \alpha \cdot m\| < \delta \Rightarrow \alpha \cdot m \in B_{\|\|}(x, \delta) \underset{\alpha \cdot m \in \alpha \cdot M}{\Rightarrow} U \bigcap (\alpha \cdot M) \neq \emptyset$ proving by 12.20 that $x \in \overline{\alpha \cdot M}$, which proves that

$$\alpha \cdot \bar{M} \subseteq \overline{\alpha \cdot M} \tag{12.12}$$

Using 12.11 and 12.12 we have that

$$\alpha \cdot \bar{M} = \overline{\alpha \cdot M}$$

2. If $x \in \alpha \cdot M$ then $x = \alpha \cdot m$ where $m \in M$. As M is open there exists a $\delta > 0$ such that $m \in B_{\|\|}(m, \delta) \subseteq M$. So if $z \in B_{\|\|}(x, |\alpha| \cdot \delta)$ then we have $\|x - z\| < |\alpha| \cdot \delta \Rightarrow \frac{1}{|\alpha|} \cdot \|x - z\| < \delta \Rightarrow \left\|\frac{1}{\alpha} \cdot x - \frac{1}{\alpha} \cdot z\right\| < m \Rightarrow \left\|m - \frac{1}{\alpha} \cdot z\right\| < \delta \Rightarrow \frac{1}{\alpha} \cdot z \in B_{\|\|}(m, \delta) \subseteq M \Rightarrow z \in \alpha \cdot M$ so that $x \in B_{\|\|}(x, |\alpha| \cdot \delta) \subseteq \alpha \cdot M$ proving that $\alpha \cdot M$ is open.

3. Let $z \in x + M$ then $z - x \in M$ which is open and there exists a $\delta > 0$ such that $(z - x) \in B_{\|\|}(z - x, \delta) \subseteq M$. If now $y \in B_{\|\|}(z, \delta)$ then $\|(y - x) - (z - x)\| = \|z - y\| < \delta$ or $y - x \in B_{\|\|}(z - x, \delta) \subseteq M \Rightarrow y \in x + M$ proving that $z \in B_{\|\|}(z, \delta) \subseteq x + M$ and thus that $x + M$ is open.

4. This follows from (3) as $A + M = \bigcup_{x \in A} (x + M)$ a union of open sets which is open by definition.

5. If $x \in \bar{M} - \bar{M}$ then $x = y_1 - y_2$, $y_1, y_2 \in \bar{M}$ and suppose that $x \in U$ a open set. So there exists a $\delta > 0$ such that $x \in B_{\|\|}(x, \delta) \subseteq U$. As $y_1, y_2 \in \bar{M}$ we have by the openness of $B_{\|\|}\left(y_1, \frac{\delta}{2}\right)$, $B_{\|\|}\left(y_2, \frac{\delta}{2}\right)$ and 12.20 that there exists $z_1 \in B_{\|\|}\left(y_1, \frac{\delta}{2}\right) \bigcap M$, $z_2 \in B_{\|\|}\left(y_2. \frac{\delta}{2}\right) \bigcap M$. Take then $z = z_1 - z_2 \in M - M$ then $\|x - z\| = \|y_1 - z_1 + y_2 - z_2\| \leqslant \|y_1 - z_1\| + \|y_2 - z_2\| < \frac{\delta}{2} + \frac{\delta}{2} = \delta$ proving that $z \in B_{\|\|}(x, \delta) \bigcap (M - M) \subseteq U \bigcap (M - M)$ so by 12.20 we have $x \in \overline{M - M}$. □

Theorem 12.77. *Let $\|\|_1$, $\|\|_2$ be two norms on a set X and $\mathcal{T}_1$, $\mathcal{T}_2$ the generated normed topologies. Then $\mathcal{T}_2$ is finer then $\mathcal{T}_1$ iff there exists a $M > 0$ such that $\forall x \in X$ we have $\|x\|_1 \leqslant M \cdot \|x\|_2$*

Proof. We use 12.59 to prove this theorem

1. **(⇒)** As $0 \in B_{\|\|_1}(0, 1)$ there exists by 12.59 a $\delta > 0$ such that $0 \in B_{\|\|_2}(0, \delta) \subseteq B_{\|\|_1}(0, 1)$ so if $\|z\|_2 < \delta \Rightarrow z \in B_{\|\|_2}(0, \delta) \subseteq B_{\|\|_1}(0, 1) \Rightarrow \|z\|_1 < 1$ so if $\|z\|_2 < \delta$ then we have $\|z\|_1 < 1$. Take now $M = \frac{2}{\delta}$ then we we have two cases to consider for $x \in X$

   a. $\boldsymbol{(x = 0)}$ then $\|x\|_1 = 0 \leqslant 0 = M \cdot 0 = M \cdot \|x\|_2$

   b. $\boldsymbol{(x \neq 0)}$ for $z = \frac{1}{M \cdot \|x\|_2} x$ we have $\|z\|_2 = \frac{1}{M \cdot \|x\|_2} \cdot \|x\|_2 = \frac{1}{M} = \frac{\delta}{2} < \delta \Rightarrow \|z\|_1 < 1 \Rightarrow \frac{1}{M \cdot \|x\|_2} \|x\|_1 < 1 \Rightarrow \|x\|_1 < M \cdot \|x\|_2 \Rightarrow \|x\|_1 \leqslant M \cdot \|x\|_2$

2. **(⇐)** If $x \in X$ and $\varepsilon > 0$ so that $x \in B_{\|\|_1}(x, \varepsilon)$ taken then $\delta = \frac{\varepsilon}{M} > 0$ (division is OK as $M > 0$ by the hypothesis). If $y \in B_{\|\|_2}(x, \delta)$ then $\|x - y\|_2 < \delta$ and we have $\|x - y\|_1 \leqslant M \cdot \|x - y\|_2 < M \cdot \delta = M \cdot \frac{\varepsilon}{M} = \varepsilon \Rightarrow y \in B_{\|\|_1}(x, \varepsilon)$ or $x \in B_{\|\|_2}(x, \delta) \subseteq B_{\|\|_1}(x, \varepsilon)$ proving by 12.59 that $\mathcal{T}_2$ is finer then $\mathcal{T}_1$ □

Lemma 12.78. *Let $A \subseteq \mathbb{R}$ then given $\alpha \in \mathbb{R}$ with $\alpha > 0$ then if $\max(A)$ exists we have that $\max(\alpha \cdot A)$ exists and $\max(\alpha \cdot A) = \alpha \cdot \max(A)$ where $\alpha \cdot A = \{\alpha \cdot a | a \in A\}$*

Proof. As $\max(A)$ exists we have that $\max(A) \in A$ and $\forall a \in A$ we have $a \leqslant \max(A)$. If now $b \in \alpha \cdot A$ then there exists a $a \in A$ such that $b = \alpha \cdot A$, from $a \leqslant \max(A)$ and $\alpha > 0$ we have that $b = \alpha \cdot a \leqslant \alpha \cdot \max(A)$. This together with the fact that $\max(A) \in A \Rightarrow \alpha \cdot \max(A) \in \alpha \cdot A$ proves that $\max(\alpha \cdot A)$ exists and is equal to $\alpha \cdot \max(A)$ □

Theorem 12.79. *Let $\{\langle X_i, \|\|_i\rangle\}_{i \in I}, I$ finite be a finite family of normed spaces then $\|\|: \prod_{i \in I} X_i \rightarrow \mathbb{R}$ defined by $\|x\| = \max(\{\|\pi_i(x)\|_i | i \in I\})$ is a norm and the generated topology on $\prod_{i \in I} X_i$ coincidences with the product topology $\mathcal{T}_{\text{product}}$ (see 12.37) (using the norm topologies of X_i). The norm $\|\|$ is called the maximum norm.*

Proof. First we prove that $\|\|$ is a norm (note that by 5.50 the maximum is always defined)

1. $\|x\| = \max(\{\|\pi_i(x)\|_i | i \in I\}) \geqslant 0$ as $\forall i \in I$ we have $\|\pi_i(x)\|_i \geqslant 0$
2. $\|\alpha \cdot x\| = \max(\{\|\pi_i(x)\|_i | i \in I\}) \underset{\pi_i(\alpha \cdot x) = (\alpha \cdot x)(i) = \alpha \cdot x(i) = \alpha \cdot \pi_i(x)}{=} \max(\{\|\alpha \cdot \pi_i(x)\|_i | i \in I\}) = \max(\{|\alpha| \cdot \|x_i\|_i | i \in I\}) \underset{12.78}{=} |\alpha| \cdot \max(\{\|x_i\|_i | i \in I\}) = |\alpha| \cdot \|x\|$
3. $\|x + y\| = \max(\{\|\pi_i(x + y)\|_i | i \in I\}) \underset{\pi_i(x+y) = (x+y)(i) = x(i) + y(i) = \pi_i(x) + \pi_i(y)}{=} \max(\{\|\pi_i(x) + \pi_i(y)\|_i | i \in I\}) \leqslant_{2.164, 5.50} \max(\{\|\pi_i(x)\|_i + \|\pi_i(y)\|_i | i \in I\}) \leqslant_{2.164} \max(\{\max(\{\pi_i(x)_i | i \in I\}) + \max(\{\|\pi_i(x)\|_i | i \in I\})\}) = \max(\{\|\pi_i(x)\|_i | i \in I\}) + \max(\{\|\pi_i(x)\|_i | i \in I\}) = \|x\| + \|y\|$
4. If $\|x\| = 0 \Rightarrow \max(\{\|\pi_i(x)\|_i | i \in I\}) = 0 \Rightarrow \forall i \in I$ we have $\|\pi_i\|_i = 0 \Rightarrow \pi_i(x) = 0 \Rightarrow \forall i \in I \vDash x(i) = 0 \Rightarrow x = 0$

Next using 12.67 we find that the product topology is generated by $d(x, y) = \max(\{d_{\|\|_i}(\pi_i(x), \pi_i(y)) | i \in I\}) = \max(\{\|\pi_i(x) - \pi_i(y)\|_i | i \in I\}) = \|x - y\|$ so $\|\|$ defines indeed the metric that generates the product topology. □

Example 12.80. The product topology of $\mathbb{R}^n = \prod_{i \in \{1, \ldots, n\}} \mathbb{R}$ is generated by the maximum norm $\|\|$ where $\|x\| = \max(\{|x_i| | i \in \{1, \ldots, n\}\})$

Definition 12.81. *Two norms on a vector space are equivalent if they generate the same topology*

Theorem 12.82. *Let X be a vector space and $\|\|_1$, $\|\|_2$ two norms on X then these norms are equivalent iff there exists $M_1, M_2 \in \mathbb{R}$ with $0 < M_1, M_2$ such that $\forall x \in X$ we have $M_1 \cdot \|x\|_1 \leqslant \|x\|_2 \leqslant \|x\|_1 \cdot M_2$*

Proof. Let $\mathcal{T}_1, \mathcal{T}_2$ be the topologies generated by the norms $\|\|_1$, $\|\|_2$

1. **(⇒)** If $\|\|_1$, $\|\|_2$ are equivalent then $\mathcal{T}_1$ is finer then $\mathcal{T}_2$ and $\mathcal{T}_2$ is finer then $\mathcal{T}_1$ so by 12.77 there exists a $M_1', M_2' > 0$ such that $\forall x \in X$we have $\|x\|_1 \leqslant \|x\|_2 \cdot M_1'$ and $\|x\|_2 \leqslant \|x\|_1 \cdot M_2' \underset{M_1' > 0}{\Rightarrow} \frac{1}{M_1'} \cdot \|x\|_1 \leqslant \|x\|_2 \leqslant \|x\|_1 \cdot M_2'$ which given $M_1 = \frac{1}{M_1'}, M_2 = M_2'$ proves that $\|x\|_1 \cdot M_1 \leqslant \|x\|_2 \leqslant \|x\|_1 \cdot M_2$
2. **(⇐)** If $\forall x \in X$ $\|x\|_1 \cdot M_1 \leqslant \|x\|_2 \leqslant \|x\|_1 \cdot M_2$ then we have $\forall x \in X$ that $\|x\|_1 \leqslant \|x\|_2 \cdot \frac{1}{M_1}$ and $\|x\|_2 \leqslant \|x\|_1 \cdot M_2$ which by 12.77 means that $\mathcal{T}_1$ is finer then $\mathcal{T}_2$ and $\mathcal{T}_2$ is finer then $\mathcal{T}_1$ so that $\mathcal{T}_1 = \mathcal{T}_2$ □

Using the above it is easy to prove that all norms in $\mathbb{R}$ or $\mathbb{C}$ are equivalent.

Example 12.83. All norms on $\mathbb{R}$ and $\mathbb{C}$ are equivalent with the norm $||$ (if we consider $\langle \mathbb{R}, +, \cdot \rangle$ a real vector space as a filed and $\langle \mathbb{C}, +, \cdot \rangle$ a complex vector space as a field).

Proof. Let $\|\|$ be a norm on $\mathbb{R}(\mathbb{C})$ then $\forall x \in \mathbb{R}$ $(\mathbb{C})$ we have $\|x\| = \|x \cdot 1\| = |x| \cdot \|1\|$ hence we have if we take $M = \|1\|$ that $M \cdot |x| = \|x\|$ or $M \cdot |x| \leqslant \|x\| \leqslant M \cdot |x|$ proving that $\|x\|$ is equivalent with $||$. □

Definition 12.84. *Let $\langle X, \|\|_X \rangle$ and $\langle Y, \|\|_Y \rangle$ be normed spaces then a isometry is a bijective function $\varphi: X \rightarrow Y$ such that $\forall x \in X$ we have that $\|x\|_X = \|\varphi(x)\|_Y$*

Theorem 12.85. *Let $\langle X, \|\|_X \rangle$ and $\langle Y, \|\|_Y \rangle$ be normed spaces and $\varphi: X - Y$ a isometry then $\varphi^{-1}: Y \rightarrow X$ is also a isometry*

Proof. First because of 2.38 we have that $\varphi^{-1}: Y \rightarrow X$ is also a bijection. Second if $y \in Y$ then $\|\varphi^{-1}(y)\|_X = \|\varphi(\varphi^{-1}(y))\|_Y = \|y\|_Y$ □

Theorem 12.86. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|Y\|_Y \rangle$ and $\langle Z, \|\|_Z \rangle$ be three normed spaces, $\varphi: X \rightarrow Y$ and $\psi: Y \rightarrow Z$ be isometries then $\psi \circ \varphi: X \rightarrow Z$ is a isometry*

Proof. By 2.46 $\psi \circ \varphi: X \rightarrow Z$ is a bijection. Further $\forall x \in X$ we have $\|(\psi \circ \varphi)(x)\|_Z = \|\psi(\varphi(x))\|_Z \underset{\psi \text{ is isometry}}{=} \|\varphi(x)\|_Y \underset{\varphi \text{ is isometry}}{=} \|x\|_X$ □

Theorem 12.87. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|Y\|_Y \rangle$ be normed spaces and $\varphi: X \rightarrow Y$ a linear isometry then if $\mathcal{T}_X, \mathcal{T}_Y$ are the topologies generated by $\|\|_X$, $\|\|_Y$ then $\mathcal{T}_X = \{\varphi^{-1}(V) | V \in \mathcal{T}_Y\} = \{U \subseteq X | \varphi(U) \in \mathcal{T}_Y\}$ and $\mathcal{T}_Y = \{V \subseteq Y | \varphi^{-1}(V) \in \mathcal{T}_X\} = \{\varphi(U) | U \in \mathcal{T}_X\}$. In other words the topologies are equivalent (see 12.64)*

Proof. The topologies $\mathcal{T}_X$, $\mathcal{T}_Y$ are generated by definition by the metrics d_X, d_Y where $d_X(x, y) = \|x - y\|_X$ and $d_Y(x', y') = \|x' - y'\|_Y$. If $x, y \in X$ then $d_Y(\varphi(x), \varphi(y)) = \|\varphi(x) - \varphi(y)\|_Y \underset{\varphi \text{ is linear}}{=} \|\varphi(x - y)\|_Y = \|x - y\|_X = d_X(x, y)$ proving that φ is also a isometry in the metric sense. Using 12.64 proves then the theorem. □

12.4 Inner product spaces

Definition 12.88. *A real inner product space $\langle X, \langle\rangle\rangle$ is a vector space X over the field $\mathbb{R}$ together with a mapping $\langle\rangle: X \times X \rightarrow \mathbb{R}$ (the inner product) satisfying the following:*

1. *$\forall x, y \in X$ we have $\langle x, y\rangle = \langle y, x\rangle$*
2. *$\forall x, y, z \in X$ and $\forall \alpha, \beta \in \mathbb{R}$ we have $\langle \alpha \cdot x + \beta \cdot y, z\rangle = \alpha \cdot \langle x, z\rangle + \beta \cdot \langle y, z\rangle$*

   ***Note* 12.89.** *Because of (1) we have also $\langle z, \alpha \cdot x + \beta \cdot y\rangle = \alpha \cdot \langle z, x\rangle + \beta \cdot \langle z, y\rangle$, so that $\langle\rangle$ is multi-linear.*

3. *$\forall x \in X$ we have $\langle x, x\rangle \geqslant 0$*
4. *$\forall x \in X$ we have $\langle x, x\rangle = 0 \Rightarrow x = 0$*

Example 12.90. Let $n \in \mathbb{N}$ then if we define $\langle\rangle: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ by $\langle (x_1, ..., x_n), (y_1, ..., y_n)\rangle = \sum_{i=1}^n x_i \cdot y_i$ then $\langle \mathbb{R}^n, \langle\rangle\rangle$ is a real inner product space

Proof.

1. $\langle (x_1, ..., x_n), (y_1, ..., y_n)\rangle = \sum_{i=1}^n x_i \cdot y_i = \sum_{i=1}^n y_i \cdot x_i = \langle (y_1, ..., y_n), (x_1, ..., x_n)\rangle$
2. $\langle \alpha \cdot (x_, ., ..., x_n) + \beta \cdot (y_1, ..., y_n), (z_1, ..., z_n)\rangle = \langle (\alpha \cdot x_1 + \beta \cdot y_1, ..., \alpha \cdot x_n + \beta \cdot y_n), (z_1, ..., z_n)\rangle = \sum_{i=1}^n (\alpha \cdot x_i + \beta \cdot y_i) \cdot z_i = \sum_{i=1}^n (\alpha \cdot x_i \cdot z_i + \beta \cdot y_i \cdot z_i) = \alpha \cdot \sum_{i=1}^n x_i \cdot z_i + \beta \cdot \sum_{i=1}^n y_i \cdot z_i = \alpha \cdot \langle (x_1, ..., x_n), (z_1, ..., z_n)\rangle + \beta \cdot \langle (y_1, ..., y_n), (z_1, ..., z_n)\rangle$
3. $\langle (x_1, ..., x_n), (x_1, ..., x_n)\rangle = \sum_{i=1}^n x_i^2 \geqslant 0$
4. If $0 = \langle (x_1, ..., x_n), (x_1, ..., x_n)\rangle = \sum_{i=1}^n x_i^2$ then if there exists a $i \in \{1, ..., n\}$ such that $x_i \neq 0$ then $0 < x_i^2$ hence $0 < x_i^2 + \sum_{j \in \{1, ..., n\} \setminus \{i\}} x_j^2 = \sum_{j \in \{i\}} x_i^2 + \sum_{j \in \{1, ..., n\} \setminus \{i\}} x_j^2 = \sum_{i=1}^n x_i^2 = 0$ a contradiction so we have $\forall i \in \{1, ..., n\}$ that $x_i = 0$ hence we have $(x_1, ..., x_n) = 0$ □

Definition 12.91. *A complex inner product space $\langle X, \langle\rangle\rangle$ is a vector space X over the field $\mathbb{C}$ together with a mapping $\langle\rangle: X \times X \rightarrow \mathbb{C}$ satisfying the following:*

1. *$\forall x, y \in X$ we have $\langle x, y\rangle = \overline{\langle y, x\rangle}$*

   ***Note* 12.92.** *Because of (1) we have $\langle x, x\rangle = \overline{\langle x, x\rangle}$ or $\langle x, x\rangle$ is real.*

2. *$\forall x, y, z \in X$ and $\alpha, \beta \in \mathbb{C}$ we have $\langle \alpha \cdot x + \beta \cdot y, z\rangle = \alpha \cdot \langle x, z\rangle + \beta \cdot \langle y, z\rangle$*

   ***Note* 12.93.** *Because of (1) we have $\langle z, \alpha \cdot x + \beta \cdot y\rangle = \overline{\langle \alpha \cdot x + \beta \cdot y, z\rangle} = \overline{\alpha \cdot \langle x, z\rangle + \beta \cdot \langle y, z\rangle} = \bar{\alpha} \cdot \langle z, x\rangle + \bar{\beta} \cdot \langle z, y\rangle$*

3. *$\forall x \in X$ we have $\langle x, x\rangle \geqslant 0$ (make sense since $\langle x, x\rangle$ is real)*
4. *$\forall x \in X$ if $\langle x, x\rangle = 0 \Rightarrow x = 0$*

Example 12.94. Let $n \in \mathbb{N}$ then if we define $\langle\rangle: \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$ by $\langle (x_1, ..., x_n), (y_1, ..., y_n)\rangle = \sum_{i=1}^n x_i \cdot \bar{y_i}$ then $\langle \mathbb{C}^n, \langle\rangle\rangle$ is a complex inner product space

Proof.

1. $\langle(x_1,\ldots,x_n),(y_1,\ldots,y_n)\rangle=\sum_{i=1}^n x_i\cdot\bar{y_i}=\sum_{i=1}^n \overline{\bar{x_i}\cdot y_i}=\overline{\sum_{i=1}^n y_i\cdot\bar{x_i}}=\overline{\langle y,x\rangle}$
2. $\langle\alpha\cdot(x,.,\ldots,x_n)+\beta\cdot(y_1,\ldots,y_n),(z_1,\ldots,z_n)\rangle=\langle(\alpha\cdot x_1+\beta\cdot y_1,\ldots,\alpha\cdot x_n+\beta\cdot y_n),(z_1,\ldots,z_n)\rangle=\sum_{i=1}^n(\alpha\cdot x_i+\beta\cdot y_i)\cdot\bar{z_i}=\sum_{i=1}^n(\alpha\cdot x_i\cdot z_i+\beta\cdot y_i\cdot\bar{z_i})=\alpha\cdot\sum_{i=1}^n x_i\cdot\bar{z_i}+\beta\cdot\sum_{i=1}^n y_i\cdot\bar{z_i}=\alpha\cdot\langle(x_1,\ldots,x_n),(z_1,\ldots,z_n)\rangle+\beta\cdot\langle(y_1,\ldots,y_n),(z_1,\ldots,z_n)\rangle$
3. $\langle(x_1,\ldots,x_n),(x_1,\ldots,x_n)\rangle=\sum_{i=1}^n x_i\cdot\bar{x_i}=\sum_{i=1}^n|x_i|^2\geqslant 0$
4. If $0=\langle(x_1,\ldots,x_n),(x_1,\ldots,x_n)\rangle=\sum_{i=1}^n|x_i|^2$ then if there exists a $i\in\{1,\ldots,n\}$ such that $x_i\neq 0$ then $0<|x_i|^2$ hence $0<|x_i|^2+\sum_{j\in\{1,\ldots,n\}\setminus\{i\}}|x_j|^2=\sum_{j\in\{i\}}|x_j|^2+\sum_{j\in\{1,\ldots,n\}\setminus\{i\}}|x_j|^2=\sum_{i=1}^n|x_i|^2=0$ a contradiction so we have $\forall i\in\{1,\ldots,n\}$ that $x_i=0$ hence we have $(x_1,\ldots,x_n)=0$ □

Lemma 12.95. *Let $\langle X,\langle\rangle\rangle$ be a real (complex) inner product space $n\in\mathbb{N}$, $y\in X$ and $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X$ then $\langle\sum_{i=1}^n x_i,y\rangle=\sum_{i=1}^n\langle x_i,y\rangle$ and $\langle y,\sum_{i=1}^n x_i\rangle=\sum_{i=1}^n\langle y,x_i\rangle$*

Proof. The first part is easely proved by induction so let $\mathcal{S}=\{n\in\mathbb{N}|\forall\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X$ we have $\langle\sum_{i=1}^n x_i,y\rangle=\sum_{i=1}^n\langle x_i,y\rangle\}$ then we have

$n=1$. this is true as for $\{x_i\}_{i\in\{1,\ldots,1\}}$ we have $\langle\sum_{i=1}^1 x_i,y\rangle=\langle x_1,y\rangle=\sum_{i=1}^1\langle x_i,y\rangle$

$n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$. let $\{x_i\}_{i\in\{1,\ldots,n+1\}}$ then we have $\langle\sum_{i=1}^{n+1}x_i,y\rangle=\langle x_{n+1}+(\sum_{i=1}^n x_i),y\rangle=\langle x_{n+1},y\rangle+\langle\sum_{i=1}^n x_i,y\rangle=\langle x_{n+1},y\rangle+\sum_{i=1}^n\langle x_i,y\rangle=\sum_{i=1}^{n+1}\langle x_i,y\rangle$ proving that $n+1\in\mathcal{S}$ □

The second part is equally easy proved by induction so let $\mathcal{S}=\{n\in\mathbb{N}|\forall\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X$ we have $\langle y,\sum_{i=1}^n x_i\rangle=\sum_{i=1}^n\langle u,x_i\rangle\}$ then we have

$n=1$. this is true as for $\{x_i\}_{i\in\{1,\ldots,1\}}$ we have $\langle y,\sum_{i=1}^1 x_i\rangle=\langle y,x_1\rangle=\sum_{i=1}^1\langle y,x_i\rangle$

$n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$. let $\{x_i\}_{i\in\{1,\ldots,n+1\}}$ then we have $\langle y,\sum_{i=1}^{n+1}x_i\rangle=\langle y,x_{n+1}+(\sum_{i=1}^n x_i)\rangle=\langle y,x_{n+1}\rangle+\langle y,\sum_{i=1}^n x_i\rangle=\langle y,x_{n+1}\rangle+\sum_{i=1}^n\langle y,x_i\rangle=\sum_{i=1}^{n+1}\langle y,x_i\rangle$ proving that $n+1\in\mathcal{S}$

Theorem 12.96. *Let $\langle X,\langle\rangle\rangle$ be a real (complex) inner product space then $\forall x\in X$ we have $\langle x,0\rangle=\langle 0,x\rangle=0$*

Proof. $\langle 0,x\rangle=\langle 0\cdot x,x\rangle=|0|\cdot\langle x,x\rangle=0.\langle x,x\rangle=0$. For the real case we have then $\langle x,0\rangle=\langle 0,x\rangle=0$ and for the complex case we have $\langle x,0\rangle=\overline{\langle 0,x\rangle}=\bar{0}=0$ □

Corollary 12.97. *Let $\langle X,\langle\rangle\rangle$ be a real (complex) inner product space then if $\forall x\in X$ we have $\langle x,y\rangle=\langle x,z\rangle$ we conclude that $y=z$*

Proof. By assumption we have $\langle z-y,y\rangle=\langle z-y,z\rangle=\langle z-y,y+(z-y)\rangle=\langle z-y,y\rangle+\langle z-y,z-y\rangle\Rightarrow\langle z-y,z-y\rangle=0\Rightarrow z-y=0\Rightarrow z=y$ □

Lemma 12.98. *If $\{s_i\}_{i\in\{1,\ldots,n\}}$ is a family of elements $\mathbb{R}$ ($\mathbb{C}_{\mathbb{R}}$) such that $\forall i\in\{1,\ldots,n\}\vDash s_i\geqslant 0$ then $\sum_{i\in\{1,\ldots,n\}} s_i=0$ implies that $\forall i\in\{1,\ldots,n\}$ we have $s_i=0$*

Proof. We prove this by induction on n. So let $S=\{n\in\mathbb{N}|\text{if } \{s_i\}_{i\in\{1,\ldots,\}} \text{ is such that } s_i\geqslant 0 \text{ then from } \sum_{i\in\{1,\ldots,n\}} s_i=0 \text{ it follows that } \forall i\in\{1,\ldots,n\} \text{ we have } s_i=0\}$, we have then

1. If $n=1$ then we have $0=\sum_{i\in\{1,\ldots,1\}} s_i=s_1\Rightarrow\forall i\in\{1,\ldots,1\}$ we have $s_i=0$ so that $1\in S$
2. Assume that $n\in S$ if now $\{s_i\}_{i\in\{1,\ldots,n+1\}}$ is such that $s_i\geqslant 0$ then if $0=\sum_{i\in\{1,\ldots,n+1\}} s_i=\sum_{i\in\{1,\ldots,n\}} s_i+s_{n+1}\underset{9.42,9.29}{\Rightarrow} s_{n+1}=0, \sum_{i\in\{1,\ldots,n\}} s_i=0\underset{n\in S}{\Rightarrow} s_{n+1}=0, \forall i\in\{1,\ldots,n\}$ we have $s_i\Rightarrow\forall i\in\{1,\ldots,n+1\}\vDash s_i\Rightarrow n+1\in S$

Using induction we have then proved the theorem. □

Theorem 12.99. *Let X be a finite dimensional real vector space, given a basis $\{e_i\}_{i\in\{1,\ldots n\}}$ define $\forall x,y\in X$ $\langle x,y\rangle=\sum_{i\in\{1,\ldots,n\}} x_i\cdot y_i$ where $x=\sum_{i\in\{1,\ldots,n\}} x_i\cdot e_i$ and $y=\sum_{i\in\{1,\ldots,n\}} y_i$ are the unique expansions of x,y in $\{e_i\}_{i\in\{1,\ldots,n\}}$. We have then that $\langle X,\langle\rangle\rangle$ is a real inner product space.*

Proof.

1. If $x,y\in X$ then $\langle x,y\rangle=\sum_{i\in\{1,\ldots,n\}} x_i\cdot y_i=\sum_{i\in\{1,\ldots,n\}} y_i\cdot x_i=\langle y,x\rangle$
2. If $x,y,z\in X$ and $\alpha,\beta\in\mathbb{R}$ then $\alpha\cdot x+\beta\cdot y=\alpha\cdot\sum_{i\in\{1,\ldots,n\}} x_i+\beta\cdot\sum_{i\in\{1,\ldots,n\}} y_i=\sum_{i\in\{1,\ldots,n\}}(\alpha_i\cdot x_i+\beta\cdot y_i)$ so that $\langle\alpha\cdot x+\beta\cdot y,z\rangle=\sum_{i\in\{1,\ldots,n\}}(\alpha\cdot x_i+\beta\cdot y_i)\cdot z_i=\sum_{i\in\{1,\ldots,n\}}(\alpha\cdot(x_i\cdot z_i)+\beta\cdot(y_i\cdot z_i))=\alpha\cdot\sum_{i\in\{1,\ldots,n\}} x_i\cdot z_i+\beta\cdot\sum_{i\in\{1,\ldots,n\}} y_i\cdot z_i=\alpha\cdot\langle x,z\rangle+\beta\cdot\langle y,z\rangle$
3. If $x\in X$ then $\langle x,x\rangle=\sum_{i\in\{1,\ldots,n\}} x_i\cdot x_i=\sum_{i\in\{1,\ldots,n\}} x_i^2\geqslant 0$ (using 9.41 and 10.62).
4. If $x\in X$ let $0=\langle x,x\rangle=\sum_{i\in\{1,\ldots,n\}} x_i^2$ then if $x\neq 0$ there exists a $i\in\{1,\ldots,n\}$ such that $x_i\neq 0$ then we have $x_i^2>0$ (see 9.41) and thus by 10.62 we have $\langle x,x\rangle=\sum_{i\in\{1,\ldots,n\}} x_i^2>0$ contradicting $\langle x,x\rangle=0$ so we must conclude that $x=0$ □

Lemma 12.100. *If $\{x_i\}_{i\in\{1,\ldots,n\}}$ is a family in $\mathbb{C}$ then $\overline{\sum_{i\in\{1,\ldots,n\}} x_i}=\sum_{i\in\{1,\ldots,n\}}\bar{x_i}$*

Proof. We prove this by induction, so let $S=\{n\in\mathbb{N}|\text{if } \{x_i\}_{i\in\{1,\ldots,n\}} \text{ is a family in } \mathbb{C} \text{ then } \overline{\sum_{i\in\{1,\ldots,n\}} x_i}=\sum_{i\in\{1,\ldots,n\}}\bar{x_i}\}$ then we have

1. If $n=0$ then $\overline{\sum_{i\in\{0,\ldots,0\}} x_i}=\bar{x_i}=\sum_{i\in\{0,\ldots,0\}}\bar{x_i}\Rightarrow 0\in S$
2. If $n\in S$ then $\overline{\sum_{i\in\{1,\ldots,n+1\}} x_i}=\overline{\sum_{i\in\{1,\ldots,n\}} x_i+x_{n+1}}=\overline{\sum_{i\in\{1,\ldots,n\}} x_i}+\overline{x_{n+1}}\underset{n\in S}{=}\sum_{i\in\{1,\ldots,n\}}\bar{x_i}+\overline{x_{n+1}}=\sum_{i\in\{1,\ldots,n+1\}}\bar{x_i}$ so that $n+1\in S$. □

Theorem 12.101. *Let X be a real (complex) finite dimensional vector space, given a basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ define $\forall x, y \in X$ $\langle x, y\rangle = \sum_{i\in\{1,\ldots,n\}} x_i \cdot \bar{y_i}$ where $x = \sum_{i\in\{1,\ldots,n\}} x_i \cdot e_i$ and $y = \sum_{i\in\{1,\ldots,0\}} y_i \cdot e_i$ are the unique expansions of x, y in $\{e_i\}_{i\in\{1,\ldots,n\}}$. We have then that $\langle X, \langle\rangle\rangle$ is a complex inner product space*

Proof.

1. If $x, y \in X$ then $\langle x, y\rangle = \sum_{i\in\{1,\ldots,n\}} x_i \cdot \bar{y_i} = \sum_{i\in\{1,\ldots,n\}} \overline{\overline{x_i \cdot \bar{y_i}}} = \sum_{i\in\{1,\ldots,n\}} \overline{\bar{x_i} \cdot \bar{\bar{y_i}}} = \overline{\sum_{i\in\{1,\ldots,n\}} \bar{x_i} \cdot y_i} = \overline{\sum_{i\in\{1,\ldots,n\}} y_i \cdot \bar{x_i}} = \overline{\langle y, x\rangle}$
2. If $x, y, z \in X$ and $\alpha, \beta \in \mathbb{C}$ then $\langle \alpha \cdot x + \beta \cdot y, z\rangle = \sum_{i\in\{1,\ldots,n\}} (\alpha \cdot x_i + \beta \cdot y_i) \cdot \bar{z_i} = \sum_{i\in\{1,\ldots,n\}} (\alpha \cdot (x_i \cdot \bar{z_i}) + \beta \cdot (y_i \cdot \bar{z_i})) = \alpha \cdot \sum_{i\in\{1,\ldots,n\}} x_i \cdot \bar{z_i} + \beta \cdot \sum_{i\in\{1,\ldots,n\}} y_i \cdot \bar{z_i} = \alpha \cdot \langle x, z\rangle + \beta \cdot \langle y, z\rangle$
3. If $x \in X$ then $\langle x, x\rangle = \sum_{i\in\{1,\ldots,n\}} x_i \cdot \bar{x_i} = \sum_{i\in\{1,\ldots,n\}} (|x_i|^2, 0) \geqslant 0$ (using 9.31 and 10.62)
4. If $x = 0$ and $\langle x, x\rangle = 0$ then $\sum_{i\in\{1,\ldots,n\}} (|x_i|^2, 0) = 0$ and using 10.62 we have $\forall i \in \{1, \ldots, n\}$ that $(|x_i|^2, 0) = 0 \Rightarrow |x_i|^2 = 0 \underset{9.31}{\Rightarrow} x_i = 0$ proving that $x = 0$. □

Definition 12.102. *Let $\langle X, \langle\rangle\rangle$ be a real (complex) inner product space then the inner product norm is defined as $\forall x \in X$ we define $\|\|: X \to \mathbb{R}$ by $\|x\| = \sqrt{\langle x, x\rangle}$*

Lemma 12.103. *Let $\langle X, \langle\rangle\rangle$ be a real (complex) inner product space then $\|0\| = 0$*

Proof. $\|0\| = \sqrt{\langle 0, 0\rangle} = \sqrt{0} = 0$ □

Theorem 12.104. (Schwartz's inequality) *If $\langle X, \langle\rangle\rangle$ is a real (complex) inner space and $\|\|$ the inner product norm then $\forall x, y \in X$ we have $|\langle x, y\rangle| \leqslant \|x\| \cdot \|y\|$. Equality holds if and only if x and y are linearly dependent.*

Proof. We have the following cases to consider

1. $(\boldsymbol{x = 0})$ then x, y are linear dependent, $\langle 0, y\rangle \underset{12.96}{=} 0$ and because of the previous lemma we have $\|x\| = \|0\| = 0$ so that we have $|\langle x, y\rangle| = |\langle 0, y\rangle| = 0 = 0 \cdot \|y\| = \|x\| \cdot \|y\|$
2. $(\boldsymbol{y = 0})$ then x, y are linear dependent, $\langle x, 0\rangle \underset{12.96}{=} 0$ and because of the previous lemma we have $\|y\| = \|0\| = 0$ so that we have $|\langle x, y\rangle| = |\langle x, 0\rangle| = 0 = \|x\| \cdot 0 = \|x\| \cdot \|y\|$
3. $(\boldsymbol{x \neq 0 \wedge y \neq 0})$ then $\langle y, y\rangle \neq 0$ so that we can define $\alpha = \frac{-\langle x, y\rangle}{\langle y, y\rangle} = -\frac{\langle x, y\rangle}{\|y\|^2}$
   a. **(real inner product)** $0 \leqslant \langle x + \alpha \cdot y, x + \alpha \cdot y\rangle = \langle x, x\rangle + 2 \cdot \alpha \cdot \langle x, y\rangle + \alpha^2 \cdot \langle y, y\rangle = \|x\|^2 - 2 \cdot \frac{\langle x, y\rangle^2}{\|y\|^2} + \frac{\langle x, y\rangle^2}{\|y\|^4} \cdot \|y\|^2 = \|x\|^2 - 2 \cdot \frac{\langle x, y\rangle^2}{\|y\|^2} + \frac{\langle x, y\rangle^2}{\|y\|^2} = \|x\|^2 - \frac{\langle x, y\rangle^2}{\|y\|^2} \underset{\text{mulitply by } \|y\|^2}{\Rightarrow} 0 \leqslant \|x\|^2 \cdot \|y\|^2 - \langle x, y\rangle^2 \Rightarrow \langle x, y\rangle^2 \leqslant \|x\|^2 \cdot \|y\|^2 = (\|x\| \cdot \|y\|)^2 \underset{9.71}{\Rightarrow} \langle x, y\rangle \leqslant \|x\| \cdot \|y\|$. Now assume that x, y are linearly dependent then we have either
      i. $(\boldsymbol{x = \beta \cdot y})$ then $\langle x, y\rangle^2 = \beta \cdot \langle x, y\rangle \cdot \langle y, y\rangle = \langle x, x\rangle \cdot \langle y, y\rangle = \|x\|^2 \cdot \|y\|^2 \underset{9.71}{\Rightarrow} |\langle x, y\rangle| = \|x\| \cdot \|y\|$

ii. $(\boldsymbol{y=\beta\cdot x})$ then $\langle x,y\rangle^2=\beta\cdot\langle x,x\rangle\cdot\langle x,y\rangle=\langle x,x\rangle\cdot\langle y,y\rangle=\|x\|^2\cdot\|y\|^2\underset{9.71}{\Rightarrow}|\langle x,y\rangle|=\|x\|\cdot\|y\|$

Assume now that $|\langle x,y\rangle|=\|x\|\cdot\|y\|$ then we have that $\langle\langle y,y\rangle\cdot x-\langle x,y\rangle\cdot y,\langle y,y\rangle\cdot x-\langle x,y\rangle\cdot y\rangle=\langle y,y\rangle^2\cdot\langle x,x\rangle-2\cdot\langle y,y\rangle\cdot\langle x,y\rangle^2+\langle x,y\rangle^2\cdot\langle y,y\rangle=\|y\|^4\cdot\|x\|^2-\|y\|^2\cdot\langle x,y\rangle^2=\|y\|^4-\|y\|^2\cdot\|x\|^2\cdot\|y\|^2=0$ proving that $\langle y,y\rangle\cdot x=\langle x,y\rangle\cdot y\Rightarrow x=\frac{\langle x,y\rangle}{\langle y,y\rangle}\cdot y$ and linear dependency

b. **(complex inner product)** $0\leqslant\langle x+\alpha\cdot y,x+\alpha\cdot y\rangle=\langle x,y\rangle+\bar{\alpha}\cdot\langle x,y\rangle+\alpha\cdot\langle y,x\rangle+\alpha\cdot\bar{\alpha}\cdot\langle y,y\rangle=\|x\|^2-\frac{\overline{\langle x,y\rangle}}{\|y\|^2}\cdot\langle x,y\rangle-\frac{\langle x,y\rangle}{\|y\|^2}\cdot\overline{\langle x,y\rangle}+|\alpha|^2\cdot\|y\|^2=\|x\|^2-2\cdot\frac{|\langle x,y\rangle|^2}{\|y\|^2}+\frac{|\langle x,y\rangle|^2}{\|y\|^4}\cdot\|y\|^2=\|x\|^2-\frac{|\langle x,y\rangle|^2}{\|y\|^2}\Rightarrow 0\leqslant\|x\|^2\cdot\|y\|^2-|\langle x,y\rangle|^2\Rightarrow|\langle x,y\rangle|^2\leqslant\|x\|^2\cdot\|y\|^2\underset{9.71}{\Rightarrow}|\langle x,y\rangle|\leqslant\|x\|\cdot\|y\|$. Now assume that x,y are linearly dependent then we have either

i. $(\boldsymbol{x=\beta\cdot y})$ then $\langle x,y\rangle^2=\beta\cdot\langle x,y\rangle\cdot\langle y,y\rangle=\langle x,x\rangle\cdot\langle y,y\rangle=\|x\|^2\cdot\|y\|^2\underset{9.71}{\Rightarrow}|\langle x,y\rangle|=\|x\|\cdot\|y\|$

ii. $(\boldsymbol{y=\beta\cdot x})$ then $\langle x,y\rangle^2=\beta\cdot\langle x,x\rangle\cdot\langle x,y\rangle=\langle x,x\rangle\cdot\langle y,y\rangle=\|x\|^2\cdot\|y\|^2\underset{9.71}{\Rightarrow}|\langle x,y\rangle|=\|x\|\cdot\|y\|$

Assume now that $|\langle x,y\rangle|=\|x\|\cdot\|y\|$ then $\langle\langle y,y\rangle\cdot x-\langle x,y\rangle\cdot y,\langle y,y\rangle\cdot x-\langle x,y\rangle\cdot y\rangle=\langle y,y\rangle^2\cdot\langle x,x\rangle-\langle y,y\rangle\cdot\overline{\langle x,y\rangle}\cdot\langle x,y\rangle-\langle x,y\rangle\cdot\langle y,y\rangle\cdot\langle y,x\rangle+\langle x,y\rangle\cdot\overline{\langle x,y\rangle}\cdot\langle y,y\rangle=\|y\|^4\cdot\|x\|^2-\|y\|^2\cdot|\langle x,y\rangle|^2-\|y\|^2\cdot|\langle x,y\rangle|^2+\|y\|^2\cdot|\langle x,y\rangle|^2=\|y\|^4\cdot\|x\|^2-2\cdot\|y\|^2\cdot\|x\|^2\cdot\|y\|^2+\|y\|^2\cdot\|x\|^2\cdot\|y\|^2=0$ meaning that $\langle y,y\rangle\cdot x=\langle x,y\rangle\cdot y\Rightarrow x=\frac{\langle x,y\rangle}{\langle y,y\rangle}\cdot y$ and linear dependency. □

Theorem 12.105. *Let $\langle X,\langle\rangle\rangle$ be a inner product space and $\|\|$ the inner product norm then $\forall x,y\in X$ we have $\|x+y\|\leqslant\|x\|+\|y\|$*

Proof. We make a distinction between the real case and the complex case

1. **(real inner product)** $\|x+y\|^2=\langle x+y,x+y\rangle=\langle x,x\rangle+2\cdot\langle x,y\rangle+\langle y,y\rangle\leqslant\|x\|^2+2\cdot|\langle x,y\rangle|+\|y\|^2\leqslant\|x\|^2+2\cdot\|x\|\cdot\|y\|+\|y\|^2=(\|x\|+\|y\|)^2\underset{9.71}{\Rightarrow}\|x+y\|\leqslant\|x\|+\|y\|$
2. **(complex inner product)** $\|x+y\|^2=\langle x+y,x+y\rangle=\langle x,x\rangle+\langle x,y\rangle+\langle y,x\rangle+\langle y,y\rangle=\|x\|^2+\langle x,y\rangle+\overline{\langle x,y\rangle}+\|y\|^2=\|x\|^2+2\cdot\mathrm{Re}(\langle x,y\rangle)+\|y^2\|\leqslant_{9.33}\|x\|^2+2\cdot|\langle x,y\rangle|+\|y\|^2\leqslant\|x\|^2+2\cdot\|x\|\cdot\|y\|+\|y\|^2=(\|x\|+\|y\|)^2\underset{9.71}{\Rightarrow}\|x+y\|\leqslant\|x\|+\|\|$ □

Theorem 12.106. *Let $\langle X,\langle\rangle\rangle$ be a inner product space and $\|\|$ the inner product norm then $\langle X,\|\|\rangle$ forms a normed space*

Proof.

1. $\|x\|=\sqrt{\langle x,x\rangle}\geqslant 0$ (or $\mathrm{Re}\left(\sqrt{\langle x,x\rangle}\right)>0$ in the complex case)

2. $\|\alpha\cdot x\| = \sqrt{(\alpha\cdot x,\alpha\cdot x)}\cdot\sqrt{\alpha^2\cdot\langle x,x\rangle} = |\alpha|\cdot\sqrt{\langle x,x\rangle} = |\alpha|\cdot\|x\|$ (or $\mathrm{Re}\left(\sqrt{\langle\alpha\cdot x,\alpha\cdot x\rangle}\right) = \mathrm{Re}\left(\sqrt{\alpha^2\cdot\langle x,x\rangle}\right) = \mathrm{Re}\left(\alpha\cdot\sqrt{\langle x,x\rangle}\right) = \alpha\cdot\mathrm{Re}\left(\sqrt{\langle x,x\rangle}\right) = \alpha\cdot\|x\|$ in the complex case)
3. Using 12.105 we have $\|x+y\|\leqslant\|x\|+\|y\|$
4. If $\|x\|=0\Rightarrow\sqrt{\langle x,x\rangle}=0\Rightarrow\langle x,x\rangle=0\Rightarrow x=0$ (or $\mathrm{Re}\left(\sqrt{\langle x,x\rangle}=0\right)\Rightarrow\sqrt{\langle x,x\rangle}=0\Rightarrow\langle x,x\rangle=0\Rightarrow x=0$ in the complex case). □

Example 12.107. Let $n\in\mathbb{N}$ then $\langle\mathbb{R}^n,\|\|_e\rangle$ where $\|(x_1,\ldots,x_n)\|=\sqrt{\sum_{i=1}^n x_i^2}$ is a normed space

Proof. This follows directly from te above theorem and 12.90. □

We can extend the above to general finite dimensional spaces

Example 12.108. If X is a finite dimensional real (complex) space with the canonical inner product defined by the basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ then $\forall x\in X$ we have $\|x\| = \sqrt{\sum_{i\in\{1,\ldots,n\}} x_i^2} = \sqrt{\sum_{i\in\{1,\ldots,n\}}|x_i|^2}$ (or $\|x\| = \mathrm{Re}\left(\sqrt{\sum_{i\in\{1,\ldots,n\}} x_i\cdot\bar{x_i}}\right) \underset{10.175}{=} \sqrt{\sum_{i\in\{1,\ldots,n\}}\mathrm{Re}(x_i\cdot\bar{x_i})} = \sqrt{\sum_{i\in\{1,\ldots,n\}}|x_i|^2}$ in the complex case) as the inner product norm (here $x=\sum_{i\in\{1,\ldots,n\}} x_i\cdot e_i$ is the unique expansion of

The following shows that $\mathbb{R}^2$ and $\mathbb{C}$ are isometric

Example 12.109. Let $\langle\mathbb{R}^2,\|\|_e\rangle$ be the normed space using the euclidean norm (see 12.107) , $\langle\mathbb{C},||\rangle$ be the normed space using the canonical norm then $\mathcal{C}:\mathbb{R}^2\rightarrow\mathbb{C}$ defined by $\mathcal{C}((x,y))=x+i\cdot y$ is a isometry

Proof. Using 10.180 we have that $\mathcal{C}$ is a isomorphims hence a bijection. Further $|\mathcal{C}((x,y))|=|x+i\cdot y|=\sqrt{x^2+y^2}=\|(x,y)\|_e$ □

In fact finite dimensional real (complex) spaces of the same dimensionality are isometric using the inner product norm.

Theorem 12.110. *Let X,Y be finite dimensional real vector spaces of the same dimensionality with basis $\{e_i\}_{i\in\{1,\ldots,n\}}$, $\{f_i\}_{i\in\{1,\ldots,n\}}$ then both spaces are isomorphic and isometric, so that the topologies generated by there inner product norm are equivalent (see 12.87)*

Proof. Define $\varphi: X\rightarrow Y$ by $x=\sum_{i\in\{1,\ldots,n\}} x_i\cdot e_i\rightarrow\varphi(x)=\sum_{i\in\{1,\ldots,n\}} x_i\cdot f_i$ then we have

1. **(injectivity)** If $\varphi(x)=\varphi(y)$ then $\sum_{i\in\{1,\ldots,n\}} x_i\cdot f_i=\sum_{i\in\{1,\ldots,n\}} y_i\cdot f_i$ then by uniqueness of the expansion in a basis we must have $\forall i\in\{1,\ldots,n\}$ that $x_i=y_i\Rightarrow x=y$
2. **(surjectivity)** If $y\in Y$ then $y=\sum_{i\in\{1,\ldots,n\}} y_i\cdot f_i$ and $\varphi\left(\sum_{i\in\{1,\ldots,n\}} y_i\cdot e_i\right)=\sum_{i\in\{1,\ldots,n\}} y_i\cdot f_i=y$

Also φ is multilinear for if $x = \sum_{i \in \{1,\ldots,n\}} x_i \cdot e_i$, $\sum_{i \in \{1,\ldots,n\}} y_i \cdot e_i$ then $\varphi(\alpha \cdot x + \beta \cdot y) = \varphi\left(\alpha \cdot \sum_{i \in \{1,\ldots,n\}} x_i \cdot e_i + \beta \cdot \sum_{i \in \{1,\ldots,n\}} y_i \cdot e_i\right) = \varphi\left(\sum_{i \in \{1,\ldots,n\}} (\alpha \cdot x_i + \beta \cdot y_i) \cdot e_i\right) = \sum_{i \in \{1,\ldots,n\}} (\alpha \cdot x_i + \beta \cdot y_i) \cdot f_i = \alpha \cdot \sum_{i \in \{1,\ldots,n\}} x_i \cdot f_i + \beta \cdot \sum_{i \in \{1,\ldots,n\}} y_i \cdot f_i = \alpha \cdot \varphi(x) + \beta \cdot \varphi(y)$

Finally we have if $x = \sum_{i \in \{1,\ldots,n\}} x_i \cdot e_i$ then $\|\varphi(x)\| = \left\|\sum_{i \in \{1,\ldots,n\}} x_i\right\| = \sqrt{\sum_{i \in \{1,\ldots,n\}} |x_i|^2} = \|x\|$ □

Definition 12.111. *Let $\langle X, \langle\rangle\rangle$ be a real (complex) inner vector space then x, $y \in X$ are **orthogonal** iff $\langle X, y\rangle = 0$*

Definition 12.112. *Let $\langle X, \langle\rangle\rangle$ be a real (complex) inner vector space then a non empty family $\{e_i\}_{i \in I}$ on X is **orthonormal** iff $\forall i, j \in I$ we have $\langle e_i, e_j\rangle = \delta_{i,j}$ (hence $\|e_i\| = 1$)*

Example 12.113. Let $n \in \mathbb{N}$ then we have that

1. Given $\langle \mathbb{R}^n, \langle\rangle\rangle$ (see 12.90) then the basis $\{\varepsilon\}_{i \in \{1,\ldots,n\}}$ (see 10.162) is orthonormal
2. Given $\langle \mathbb{C}^n, \langle\rangle\rangle$ (see 12.94) then the basis $\{\varepsilon\}_{i \in \{1,\ldots,n\}}$ (see 10.162) is orthonormal

Proof.

1. Given $i, j \in \{1, \ldots, n\}$ we have that $\langle \varepsilon_i, \varepsilon_j\rangle = \sum_{k=1}^{n} (\varepsilon_i)_k \cdot (\varepsilon_j)_k = \sum_{k=1}^{n} \delta_{i,k} \cdot \delta_{j,k} = \delta_{i,j}$
2. Given $i, j \in \{1, \ldots, n\}$ we have that $\langle \varepsilon_i, \varepsilon_j\rangle = \sum_{k=1}^{n} (\varepsilon_i)_k \cdot \overline{(\varepsilon_j)_k} = \sum_{k=1}^{n} \delta_{i,k} \cdot \overline{\delta_{j,k}} = \delta_{i,j}$ □

Every orthonormal family is linear independent (the opposite is of course not always true)

Theorem 12.114. *Let $\langle X, \langle\rangle\rangle$ be a real (complex) inner product vector space then every orthonormal family $\{e_i\}_{i \in I}$ is linear independent.*

Proof. Let $\{a_i\}_{i \in I}$ is a family in $\mathbb{R}(\mathbb{C})$ with finite support such that $0 = \sum_{i \in I} \alpha_i \cdot e_i$ then $\forall j \in I$ we have that $0 = \langle 0, e_j\rangle = \langle \sum_{i \in I} \alpha_i \cdot e_i, e_j\rangle$. Now as $\{a_i\}_{i \in I}$ has finite support and thus $\{\alpha_i \cdot e_i\}_{i \in I}$ has finite support there exists a $n \in \mathbb{N}_0$ and a bijection $\varphi: \{0, \ldots, n\} \rightarrow I$ such that $\sum_{i \in I} \alpha_i \cdot e_i = \sum_{i=0}^{n} \alpha_{\varphi(i)} \cdot e_{\varphi(i)}$ hence $0 = \langle \sum_{i \in I} \alpha_i \cdot e_i, e_j\rangle = \langle \sum_{i=0}^{n} \alpha_{\varphi(i)} \cdot e_{\varphi(i)}, e_j\rangle \underset{12.95}{=} \sum_{i=0}^{n} \langle \alpha_{\varphi(i)} \cdot e_{\varphi(i)}, e_j\rangle = \sum_{i=0}^{n} \alpha_{\varphi(i)} \cdot \delta_{\varphi(i),j} = \sum_{i \in I} \alpha_i \cdot \delta_{i,j} = \sum_{i \in I \setminus \{j\}} \alpha_i \cdot \delta_{i,j} + \sum_{i \in \{j\}} \alpha_i \cdot \delta_{i,j} = 0 + \alpha_j = \alpha_j$ proving that $\forall j \in I$ we have $\alpha_j = 0$. □

Theorem 12.115. *Let $\langle X, \langle\rangle\rangle$ be a real (complex) inner product space with a orthonormal basis $\{e_i\}_{i \in \{1,\ldots,n\}}$, $L \in \mathrm{Hom}(X)$ then $\forall i, j \in \{1, \ldots, n\}$ $\mathcal{M}(L, \{e_i\}_{i \in \{1,\ldots,n\}})_{j,i} = \langle L(e_i), e_j\rangle$*

Proof.

$$\begin{aligned}\langle L(e_i), e_j\rangle &= \left\langle \sum_{k=1}^{n} \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}})_{k,i} \cdot e_k, e_j \right\rangle \\ &= \sum_{k=1}^{n} \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}})_{k,i} \cdot \langle e_k, e_j\rangle \\ &= \sum_{k=1}^{n} \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}})_{k,i} \cdot \delta_{k,j} \\ &= \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}})_{j,i}\end{aligned}$$

□

It turns out that every finite dimensional real (complex) inner vector space has a orthonormal basis, for this we use the Gram-Schmidt procedure.

Theorem 12.116. *Let $\langle X, \langle\rangle\rangle$ be a real (complex) inner vector space and $\{e_i\}_{i\in\{1,\ldots,n\}}$, $n\in\mathbb{N}$ a linear independent family of vectors in X then if we define $\{f_i\}_{i\in\{1,\ldots,n\}}$ by $f_1 = \frac{e_1}{\|e_1\|}$ and $\forall i\in\{2,\ldots,n\}$ $\hat{e}_i = \frac{e_i - \sum_{j=1}^{i-1}\langle e_i, \hat{e}_j\rangle\cdot\hat{e}_j}{\|e_i - \sum_{j=1}^{i-1}\langle e_i, \hat{e}_j\rangle\cdot\hat{e}_j\|}$ then $\{\hat{e}_i\}_{i\in\{1,\ldots,n\}}$ is well defined (no division by 0), orthonormal and $\mathcal{S}(\{e_i|i\in\{1,\ldots,j\}\}) = \mathcal{S}(\{\hat{e}_i|i\in\{1,\ldots,j\}\})$ for $j\in\{1,\ldots,n\}$ (the spans are equal).*

Proof. We will use induction to prove this so let $\mathcal{F} = \{n\in\mathbb{N}|\text{if } \{e_i\}_{i\in\{1,\ldots,n\}}$ then $\{\hat{e_i}\}_{i\in\{1,\ldots,n\}}$ is well defined, orthonormal and $j\in\{1,\ldots,n\}$ $\mathcal{S}(\{e_i|i\in\{1,\ldots,j\}\}) = \mathcal{S}(\{\hat{e}_i|i\in\{1,\ldots,j\}\})\}$ then we have

$1\in\mathcal{F}$. As $\{e_i\}_{i\in\{1,\ldots,1\}}$ is linear independent we have that $e_1\neq 0 \Rightarrow \|e_1\|\neq 0$ so $\hat{e}_1 = \frac{e_1}{\|e_1\|}$ is well defined $\langle\hat{e}_1, \hat{e}_1\rangle = \left\langle\frac{e_1}{\|e_1\|}, \frac{e_1}{\|e_i\|}\right\rangle = \frac{\langle e_1, e_1\rangle}{\|e_1\|^2} = \frac{\|e_1\|^2}{\|e_1\|^2} = 1$ and $\mathcal{S}(\{e_1\}) = \{\alpha\cdot e_1|\alpha\in\mathbb{R}(\mathbb{C})\} = \left\{\alpha\cdot\frac{e_1}{\|e_1\|}|\alpha\in\mathbb{R}(\mathbb{C})\right\} = \mathcal{S}(\{\hat{e}_1\})$ proving that $1\in\mathcal{F}$

$n\in\mathcal{F}\Rightarrow n+1\in\mathcal{F}$. Let $\{e_i\}_{i\in\{1,\ldots,n+1\}}$ be a linear independent set, then as $n\in\mathcal{F}$ we have that

$$\mathcal{S}(\{e_i|i\in\{1,\ldots,j\}|\}) = \mathcal{S}(\{\hat{e}_i|i\in\{1,\ldots,j\}\})\,\forall j\in\{1,\ldots,n\} \tag{12.13}$$

Now as $\{e_i\}_{i\in\{1,\ldots,n+1\}}$ is linear independent we have that $e_{n+1}\notin\mathcal{S}(\{e_i|i\in\{1,\ldots,n\}|\}) = \mathcal{S}(\{\hat{e}_i|i\in\{1,\ldots,n\}\})$ so that $e_{n+1}\neq\sum_{k=1}^{n}\langle e_i, \hat{e}_k\rangle\cdot\hat{e}_k$ and thus that $\|e_{n+1} - \sum_{k=1}^{n}\langle e_{n+1}, \hat{e}_k\rangle\cdot\hat{e}_k\|$ so that $\hat{e}_{n+1}$ is well defined (no division by zero). Further from the definition of $\hat{e}_{n+1}$ it follows that $e_{n+1} = \|e_{n+1} - \sum_{k=1}^{n+1}\langle e_{n+1}, \hat{e}_k\rangle\cdot\hat{e}_k\|\cdot\hat{e}_{n+1} + \sum_{k=1}^{n}\langle e_{n+1}, \hat{e}_k\rangle\cdot\hat{e}_k$ proving that $e_{n+1}\in\mathcal{S}(\{\hat{e}_i|i\in\{1,\ldots,n+1\}\})$ hence as $\{e_i|i\in\{1,\ldots,n\}\}\subseteq\mathcal{S}(\{e_i|i\in\{1,\ldots,n\}|\}) = \mathcal{S}(\{\hat{e}_i|i\in\{1,\ldots,n\}\})$ we have that $\{e_i|i\in\{1,\ldots,n+1\}\}\subseteq\mathcal{S}(\{\hat{e}_i|i\in\{1,\ldots,n\}\})\subseteq\mathcal{S}(\{\hat{e}_i|i\in\{1,\ldots,n+1\}\})$ we have that $\mathcal{S}(\{e_i|i\in\{1,\ldots,n+1\}\})\subseteq\mathcal{S}(\mathcal{S}(\{\hat{e}_i|i\in\{1,\ldots,n+1\}\})) \underset{10.135}{=} \mathcal{S}(\{\hat{e}_i|i\in\{1,\ldots,n+1\}\})$ proving

$$\mathcal{S}(\{e_i|i\in\{1,\ldots,n+1\}|\})\subseteq\mathcal{S}(\{\hat{e}_i|i\in\{1,\ldots,n+1\}\}) \tag{12.14}$$

Next for $i, j\in\{1,\ldots,n+1\}$ we have to consider the following cases

$i, j\in\{1,\ldots,n\}$. then as $n\in\mathcal{F}$ we have $\langle\hat{e}_i, \hat{e}_j\rangle = \delta_{i,j}$

$\boldsymbol{i=n+1, j\in\{1,\ldots,n\}}$. then

$$\begin{aligned}\langle\hat{e}_i,\hat{e}_j\rangle &= \left\langle\frac{e_i-\sum_{k=1}^{i-1}\langle e_i,\hat{e}_k\rangle\cdot\hat{e}_k}{\|e_i-\sum_{k=1}^{i-1}\langle e_i,\hat{e}_k\rangle\cdot\hat{e}_k\|},\hat{e}_j\right\rangle\\ &= \frac{1}{\|e_i-\sum_{k=1}^{i-1}\langle e_i,\hat{e}_k\rangle\cdot\hat{e}_k\|}\cdot\left(\langle e_i,\hat{e}_j\rangle-\sum_{k=1}^{i-1}\langle e_i,\hat{e}_k\rangle\cdot\langle\hat{e}_k,\right.\\ &\left.\hat{e}_j\rangle\right)\\ &\underset{n\in\mathcal{F}}{=} \frac{1}{\|e_i-\sum_{k=1}^{i-1}\langle e_i,\hat{e}_k\rangle\cdot\hat{e}_k\|}\cdot\left(\langle e_i,\hat{e}_j\rangle-\sum_{k=1}^{i-1}\langle e_i,\hat{e}_k\rangle\cdot\delta_{k,j}\right)\\ &= \frac{1}{\|e_i-\sum_{k=1}^{i-1}\langle e_i,\hat{e}_k\rangle\cdot\hat{e}_k\|}\cdot(\langle e_i,\hat{e}_j\rangle-\langle e_i,\hat{e}_j\rangle)\\ &= 0\end{aligned}$$

$\boldsymbol{i\in\{1,\ldots,n\}\wedge j=n+1}$. then $\langle\hat{e}_i,\hat{e}_j\rangle=\overline{\langle\hat{e}_i,\hat{e}_j\rangle}\underset{\text{previous case}}{=}0$

$\boldsymbol{i=j=n+1}$. then

$$\begin{aligned}\langle\hat{e}_i,\hat{e}_j\rangle &= \langle\hat{e}_{n+1},\hat{e}_{n+1}\rangle\\ &= \left\langle\frac{e_{n+1}-\sum_{k=1}^{n}\langle e_i,\hat{e}_k\rangle\cdot\hat{e}_k}{\|e_{n+1}-\sum_{k=1}^{n}\langle e_i,\hat{e}_k\rangle\cdot\hat{e}_k\|},\right.\\ &\left.\frac{e_{n+1}-\sum_{k=1}^{n}\langle e_{n+1},\hat{e}_k\rangle\cdot\hat{e}_k}{\|e_{n+1}-\sum_{k=1}^{n+1}\langle e_{n+1},\hat{e}_k\rangle\cdot\hat{e}_k\|}\right\rangle\\ &= \frac{1}{\|e_{n+1}-\sum_{k=1}^{n}\langle e_{n+1},\hat{e}_k\rangle\cdot\hat{e}_k\|^2}\cdot\left\langle e_{n+1}-\sum_{k=1}^{n}\langle e_{n+1},\hat{e}_k\rangle\cdot\right.\\ &\left.\hat{e}_k,e_{n+1}-\sum_{k=1}^{n}\langle e_{n+1},\hat{e}_k\rangle\cdot\hat{e}_k\right\rangle\\ &= \frac{\|e_{n+1}-\sum_{k=1}^{i-1}\langle e_{n+1},\hat{e}_k\rangle\cdot\hat{e}_k\|^2}{\|e_{n+1}-\sum_{k=1}^{i-1}\langle e_{n+1},\hat{e}_k\rangle\cdot\hat{e}_k\|^2}\\ &= 1\end{aligned}$$

Proving that $\{\hat{e}_i\}_{i\in\{1,\ldots,n+1\}}$ is orthonormal. As we have just proved that $\{\hat{e}_i\}_{i\in\{1,\ldots,n+1\}}$ is orthonormal we have by 12.114 that $\{\hat{e}_i\}_{i\in\{1,\ldots,n+1\}}$ is a linear independent set and as $\mathcal{S}(\{\hat{e}_i|i\in\{1,\ldots,n+1\}\})$ is a vector space (see 10.130) we have that $\mathcal{S}(\{\hat{e}_i|i\in\{1,\ldots,n+1\}\})$ is a $n+1$ dimensional vector space. As by 12.14 we have that $\{e_i|i\in\{1,\ldots,n+1\}|\}$ is a linear independent subset of $\mathcal{S}(\{\hat{e}_i|i\in\{1,\ldots,n+1\}\})$ we have using 10.207 that $\{e_i\}_{i\in\{1,\ldots,n+1\}}$ is a basis for $\mathcal{S}(\{\hat{e}_i|i\in\{1,\ldots,n+1\}\})$ hence we have $\mathcal{S}(\{e_i|i\in\{1,\ldots,n+1\}\})=\mathcal{S}(\{\hat{e}_i|i\in\{1,\ldots,n+1\}\})$ and using 12.13 we conclude that

$$\forall j\in\{1,\ldots,n+1\}\,\mathcal{S}(\{e_i|i\in\{1,\ldots,n+1\}\})=\mathcal{S}(\{\hat{e}_i|i\in\{1,\ldots,n+1\}\})\tag{12.15}$$

which finally proves that $n+1 \in \mathcal{F}$ $\square$

Corollary 12.117. *Let $\langle X, \langle\rangle\rangle$ be a finite dimensional inner vector space with dimension $n \in \mathbb{N}$ (so a non trivial vector space) then we have a orthonormal basis $\{e_i\}_{i \in \{1,\ldots,n\}}$ on X*

Proof. As X is a n-dimensional vector space then there exists a linear independent family $\{f_i\}_{i \in \{1,\ldots,n\}}$ (the beasis of X) such that $\mathcal{S}(\{f_i | i \in \{1,\ldots,n\}\}) = X$, using the Gram-Schmidt procedure we find a orthonormal family (linear independent by 12.114) $\{e_i\}_{i \in \{1,\ldots,n\}}$ such that $S(\{e_i | i \in \{1,\ldots,n\}\}) = \mathcal{S}(\{f_i | i \in \{1,\ldots,n\}\}) = X$. So $\{e_i\}_{i \in \{1,\ldots,n\}}$ is a orthonormal basis of X. $\square$

Theorem 12.118. *Let $\langle X, \langle\rangle\rangle$ be a finite dimensional inner vector space with dimension $n \in \mathbb{N}$ and $\{e_i\}_{i \in \{1,\ldots,n\}}$ a orthonormal family then $\{e_i\}_{i \in \{1,\ldots,n\}}$ is a basis of X*

Proof. Let $\{\alpha_i\}_{i \in \{1,\ldots,n\}} \subseteq \mathbb{R}(\mathbb{C})$ such that $\sum_{i=1}^{n} \alpha_i \cdot e_i = 0$ then $\forall j \in \{1,\ldots,n\}$ we have $0 = \langle \sum_{i=1}^{n} \alpha_i \cdot e_i, e_j \rangle = \sum_{i=1}^{n} \alpha_i \cdot \langle e_i, e_j \rangle = \sum_{i=1}^{n} \alpha_i \cdot \delta_{i,j} = \alpha_j$, hence we have that $\{e_i\}_{i \in \{1,\ldots,n\}}$ is linear independent. Using 10.203 proves then that $\{e_i\}_{i \in \{1,\ldots,n\}}$ forms a basis of X. $\square$

Theorem 12.119. (Riesz Representation Theorem) *Let $\langle X, \langle\rangle\rangle$ be a finite dimensional real (complex) inner product space and $\varphi \in \mathrm{Hom}(X, \mathbb{R})$ (or $\mathrm{Home}(X, \mathbb{C})$) then there exists a* ***unique*** *$u \in X$ such that $\forall x \in X$ we have $\varphi(x) = \langle x, u \rangle$*

Proof. Using 12.117) we have that there exists a orthormal basis $\{e_i\}_{i \in \{1,\ldots,n\}}$ of X. Hence if $x \in X$ there exists a unique $\{x_i\}_{i \in \{1,\ldots,n\}}$ such that $x = \sum_{i=1}^{n} x_i \cdot e_i$. Then $\varphi(x) = \varphi(\sum_{i=1}^{n} x_i \cdot e_i) = \sum_{i=1}^{n} x_i \cdot \varphi(e_i)$. Take then $u = \sum_{i=1}^{n} \varphi(e_i) \cdot e_i$ then

$$\begin{aligned}
\langle x, u \rangle &= \left\langle x, \sum_{i=1}^{n} \varphi(e_i) \cdot e_i \right\rangle \\
&= \sum_{i=1}^{n} \varphi(e_i) \cdot \langle x, e_i \rangle \\
&= \sum_{i=1}^{n} \varphi(e_i) \cdot \left\langle \sum_{k=1}^{n} x_k \cdot e_k, e_i \right\rangle \\
&= \sum_{i=1}^{n} \varphi(e_i) \cdot \sum_{k=1}^{n} x_k \cdot \langle e_k, e_i \rangle \\
&= \sum_{i=1}^{n} \varphi(e_i) \cdot (\sum_{k=1}^{n} x_k \cdot \delta_{k,i} \\
&= \sum_{i=1}^{n} x_i \cdot \varphi(e_i) \\
&= \varphi(x)
\end{aligned}$$

proving the theorem. $\square$

Riesz Representation Theorem allows us to define the adjoint of a operator

Theorem 12.120. *Let $\langle X, \langle\rangle_X\rangle$, $\langle Y, \langle\rangle_Y\rangle$ be finite dimensional real (complex) inner product spaces and $L \in \mathrm{Hom}(X, Y)$ then $L^*: Y \to X$ is defined by $L^*(y)$ is the unique number such that $\langle L(x), y\rangle_Y = \langle x, L^*(y)\rangle_X$ $\forall x \in X$. L^* is called the adjoint of L.*

Proof. First we have to prove that given $y \in Y$ such a number $L^*(y)$ exists. To do this given $y \in Y$ define $\varphi_y: X \to \mathbb{R}(\mathbb{C})$ by $\varphi_y(x) = \langle L(x), y\rangle_Y$. For $x_1, x_2 \in X, \alpha$, $\beta \in \mathbb{R}(\mathbb{C})$ we have then that $\varphi_y(\alpha \cdot x_1 + \beta \cdot y_1) = \langle L(\alpha \cdot x_1 + \beta \cdot y_1), y\rangle_Y = \langle \alpha \cdot L(x_1) + \beta \cdot L(x_2), y\rangle_Y = \alpha \cdot \langle L(x_1), y\rangle_Y + \beta \cdot \langle L(x_2), y\rangle_Y = \alpha \cdot \varphi_y(x_1) + \beta \cdot \varphi_y(x_2)$ proving that $\varphi_y \in \mathrm{Hom}(X, \mathbb{R})$(or $\mathrm{Hom}(Y, \mathbb{C})$). Using the Reisz Representation Theorem (see 12.119) there exists then a **unique** $L^*(y) \in X$ such that $\forall x \in X$ such that $\forall x \in X$ we have $\varphi_y(x) = \langle x, L^*(y)\rangle_X$ or using the definition of $\varphi_y(x)$ we have $\langle x, L^*(y)\rangle_X = \langle L(x), y\rangle_X$ $\square$

Theorem 12.121. *Let $\langle X, \langle\rangle_X\rangle$, $\langle Y, \langle\rangle_Y\rangle$ be finite dimensional real (complex) inner product spaces and $L \in \mathrm{Hom}(X, Y)$ then $(L^*)^*: X \to Y$ is equal to $L: X \to Y$*

Proof. As $L: X \to Y$ we have $L^*: Y \to X$ and thus $(L^*)^*: X \to Y$. Let $x \in X$, $y \in Y$ then we have two cases to consider

***X* is a real space.**

$$\begin{aligned}\langle y, (L^*)^*(x)\rangle_Y &= \langle L^*(y), x\rangle_X \\ &= \langle x, L^*(y)\rangle_X \\ &= \langle L(x), y\rangle_Y \\ &= \langle y, L(x)\rangle_X\end{aligned}$$

***X* is a complex space.**

$$\begin{aligned}\langle y, (L^*)^*(x)\rangle_Y &= \langle L^*(y), x\rangle_X \\ &= \overline{\langle x, L^*(y)\rangle_X} \\ &= \overline{\langle L(x), y\rangle_Y} \\ &= \langle y, L(x)\rangle_Y\end{aligned}$$

So we have in all cases that $\langle y, (L^*)^*(x)\rangle_Y = \langle y, L(x)\rangle_Y$ proving by definition of the adjoint that $(L^*)^* = L$ $\square$

The operator $*$ is linear as proved in the following operator

Theorem 12.122. *Let $\langle X, \langle\rangle_X\rangle$, $\langle Y, \langle\rangle_Y\rangle$ be finite dimensional real (complex) inner product spaces, $L_1, L_2 \in \mathrm{Hom}(X, Y)$ and $\alpha, \beta \in \mathbb{R}$ then $(\alpha \cdot L_1 + \beta \cdot L_2)^* = \alpha \cdot L_1^* + \beta \cdot L_2^*$*

Proof. Given $x \in X$, $y \in Y$ then C

$$\begin{array}{rcl}\langle x, (\alpha \cdot L_1 + \beta \cdot L_2)^*(y)\rangle & = & \langle(\alpha \cdot L_1 + \beta \cdot L_2)(x), y\rangle \\ & = & \langle \alpha \cdot L_1(x) + \beta \cdot L_2(x), y\rangle \\ & = & \alpha \cdot \langle L_1(x), y\rangle + \beta \cdot \langle L_2(x), y\rangle \\ & = & \alpha \cdot \langle x, L_1^*(y)\rangle + \beta \cdot \langle x, L_2^*(y)\rangle \\ & \underset{\alpha,\beta \in \mathbb{R}}{=} & \langle x, \alpha \cdot L_1^*(y)\rangle + \langle x, \beta \cdot L_2^*(y)\rangle \\ & = & \langle x, \alpha \cdot L_1^*(y) + \beta \cdot L_2^*(y)\rangle \\ & = & \langle x, (\alpha \cdot L_1^* + \beta \cdot L_1^*)(y)\rangle\end{array}$$

proving by the definition of the adjoint that $(\alpha \cdot L_1 + \beta \cdot L_2)^* = \alpha \cdot L_1^* + \beta \cdot L_2^*$ □

Theorem 12.123. *Let $\langle X, \langle\rangle_X\rangle$, $\langle Y, \langle\rangle_Y\rangle$ and $\langle Z, \langle\rangle_Z\rangle$ be finite dimensional real (complex) inner product spaces and $L_1 \in \mathrm{Hom}(X, Y)$, $L_2 \in \mathrm{Hom}(Y, Z)$ then $(L_2 \circ L_1): X \to Z$ has a adjoint $(L_2 \circ L_1)^*: Z \to X$ defined by $(L_2 \circ L_1)^* = L_1^* \circ L_2^*$*

Proof. Let $x \in X$, $z \in Z$ then

$$\begin{array}{rcl}\langle x, (L_2 \circ L_1)^*(z)\rangle_X & = & \langle (L_2 \circ L_1)(x), z\rangle_Z \\ & = & \langle L_2(L_1(x)), z\rangle_Z \\ & = & \langle L_1(x), L_2^*(z)\rangle_Y \\ & = & \langle x, L_1^*(L_2^*(z))\rangle_X \\ & = & \langle x, (L_1^* \circ L_2^*)(z)\rangle\end{array}$$

which by definition means that $(L_2 \circ L_1)^* = L_1^* \circ L_2^*$ □

To see the relation between the adjoint of a linear operator and itst matrix we define the concept of the transpose of a matrix and the conjugate transpose of a matrix.

Definition 12.124. *Let $n, m \in \mathbb{N}$, $\mathcal{F}$ a field and $M \in \mathcal{M}(n \times m, \mathcal{F})$ then M^T is defined by $(M^T)_{i,j} = M_{j,i}$ $\forall i, j \in \{1, ..., n\} \times \{1, ..., m\}$ (see 10.296)*

Definition 12.125. *Let $n, m \in \mathbb{N}$, $\mathcal{F} = \mathbb{R}$ or $\mathcal{F} = \mathbb{C}$, $M \in \mathcal{M}(n \times m, \mathcal{F})$ then M^H is defined by $(M^H)_{i,j} = \overline{M_{j,i}}$ $\forall i, j \in \{1, ..., n\} \times \{1, ..., m\}$. Note that in the case $\mathcal{F} = \mathbb{R}$ we have that $M^H = M^T$*

Definition 12.126. (Symmetric Matrix) *Let $n, m \in \mathbb{N}$, $\mathcal{F}$ a field then $M \in \mathcal{M}(n \times m, \mathcal{F})$ is symmetric if $M^T = M$*

Definition 12.127. (Hermitian Matrix) *Let $n, m \in \mathbb{N}$, $\mathcal{F}$ a field then $M \in \mathcal{M}(n \times m, \mathcal{F})$ is hermitian if $M^H = M$. Note that in the real case a symmetric matrix is Hermitian.*

Definition 12.128. *Let $n \in \mathbb{N}$, $\mathcal{F} = \mathbb{R}$ (or $\mathbb{C}$) and $M \in \mathcal{M}(n \times b, \mathcal{F})$ then M is **unitary** if $U^H \cdot U = E = U \cdot U^H$*

Theorem 12.129. *Let $\langle X, \langle\rangle_X\rangle$, $\langle Y, \langle\rangle_Y\rangle$ be finite dimensional real (complex) inner product spaces and $L \in \operatorname{Hom}(X, Y)$ then $L^*: Y \rightarrow X$ is linear hence $L^* \in \operatorname{Hom}(Y, X)$. Further if $\{e_i\}_{i\in\{1,\ldots,n\}}$ is a orthonormal basis of X and $\{f_i\}_{i\in\{1,\ldots,m\}}$ is a orthonormal basis of Y then $\forall (i, j) \in \{1, \ldots, n\} \times \{1, \ldots, m\}$ we have $\mathcal{M}(L^*, \{f_k\}_{k\in\{1,\ldots,m\}}, \{e_k\}_{k\in\{1,\ldots,n\}})_{i,j} = \mathcal{M}(L, \{e_k\}_{k\in\{1,\ldots,n\}}, \{f_k\}_{k\in\{1,\ldots,n\}})^H \underset{\text{definition}}{=} \overline{\mathcal{M}(L, \{e_k\}_{k\in\{1,\ldots,n\}}, \{f_k\}_{k\in\{1,\ldots,m\}})_{j,i}}$ in the complex case and $\mathcal{M}(L^*, \{f_k\}_{k\in\{1,\ldots,m\}}, \{e_k\}_{k\in\{1,\ldots,n\}})_{i,j} = \mathcal{M}(L, \{e_k\}_{k\in\{1,\ldots,n\}}, \{f_k\}_{k\in\{1,\ldots,m\}})_{j,i} =$ in the real case.*

Proof. We have to make a distinction between the real and complex case.

Real Case. Then we have that $\forall x \in X, \forall \alpha, \beta \in \mathbb{R}$ and $y_1, y_2 \in Y$

$$\begin{aligned} \langle x, \alpha \cdot L^*(y_1) + \beta \cdot L^*(y_2)\rangle_X &= \alpha \cdot \langle x, L^*(y_1)\rangle + \beta \cdot \langle x, L^*(y_2)\rangle_X \\ &= \alpha \cdot \langle L(x), y_1\rangle_Y + \beta \cdot \langle L(x), y_2\rangle_Y \\ &= \langle L(x), \alpha \cdot y_1 + \beta \cdot y_2\rangle_Y \\ &= \langle x, L^*(\alpha \cdot y_1 + \beta \cdot y_2)\rangle_X \end{aligned}$$

hence using 12.97 we have that $L^*(y_1) + \beta \cdot L^*(y_2) = L^*(\alpha \cdot y_1 + \beta \cdot y_2)$ which proves linearity. Take now $M = \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,n\}})$ and $M^* = \mathcal{M}(L^*, \{f_i\}_{i\in\{1,\ldots,m\}}, \{e_i\}_{i\in\{1,\ldots,n\}})$ then by definition (see 10.300) we have for $i \in \{1, \ldots, n\}, j \in \{1, \ldots, m\}$ that

$$\begin{aligned} \langle e_i, L^*(f_j)\rangle &\underset{10.300}{=} \left\langle e_i, \sum_{k=1}^{n} M^*_{k,j} \cdot e_k \right\rangle \\ &= \sum_{k=1}^{n} M^*_{k,j} \langle e_i, e_k\rangle \\ &= \sum_{k=1}^{n} M^*_{k,j} \cdot \delta_{i,k} \\ &= M^*_{i,j} \end{aligned}$$

Using the above we have then that

$$\begin{aligned} M^*_{i,j} &= \langle e_i, L^*(f_j)\rangle \\ &= \langle L(e_i), f_j\rangle \\ &\underset{10.300}{=} \left\langle \sum_{j=1}^{m} M_{k,i} \cdot f_k, f_j \right\rangle \\ &= \sum_{j=1}^{m} M_{k,i} \cdot \langle f_k, f_j\rangle \\ &= \sum_{j=1}^{m} M_{k,i} \cdot \delta_{k,j} \\ &= M_{j,i} \end{aligned}$$

proving that $\mathcal{M}(L,\{e_k\}_{k\in\{1,\ldots,n\}},\{f_k\}_{k\in\{1,\ldots,m\}})_{i,j}=\mathcal{M}(L^*,\{f_k\}_{k\in\{1,\ldots,m\}},\{e_k\}_{k\in\{1,\ldots,n\}})_{j,i}$ where $(i,j)\in\{1,\ldots,n\}\times\{1,\ldots,m\}$

Complex Case. Then we have that $\forall x\in X,\forall\alpha,\beta\in\mathbb{C}$ and $y_1,y_2\in Y$

$$\begin{aligned}\langle x,\alpha\cdot L^*(y_1)+\beta\cdot L^*(y_2)\rangle_X &= \bar{\alpha}\cdot\langle x,L^*(y_1)\rangle+\bar{\beta}\cdot\langle x,L^*(y_2)\rangle_X\\ &= \bar{\alpha}\cdot\langle L(x),y_1\rangle_Y+\bar{\beta}\cdot\langle L(x),y_2\rangle_Y\\ &= \langle L(x),\alpha\cdot y_1+\beta\cdot y_2\rangle_Y\\ &= \langle x,L^*(\alpha\cdot y_1+\beta\cdot y_2)\rangle_X\end{aligned}$$

hence using 12.97 we have that $L^*(y_1)+\beta\cdot L^*(y_2)=L^*(\alpha\cdot y_1+\beta\cdot y_2)$ which proves linearity. Take now $M=\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,n\}})$ and $M^*=\mathcal{M}(L^*,\{f_i\}_{i\in\{1,\ldots,m\}},\{e_i\}_{i\in\{1,\ldots,n\}})$ then by definition (see 10.300) we have for $i\in\{1,\ldots,n\}, j\in\{1,\ldots,m\}$ that

$$\begin{aligned}\langle L^*(f_j),e_i\rangle &\underset{10.300}{=} \left\langle\sum_{k=1}^{m}M^*_{k,j}\cdot e_k,e_i\right\rangle\\ &= \sum_{k=1}^{m}M^*_{k,j}\langle e_k,e_i\rangle\\ &= \sum_{k=1}^{m}M^*_{k,j}\cdot\delta_{k,i}\\ &= M^*_{i,j}\end{aligned}$$

Using the above we have then that

$$\begin{aligned}M^*_{i,j} &= \langle L^*(f_j),e_i\rangle\\ &= \overline{\langle e_i,L^*(f_j)\rangle}\\ &= \overline{\langle L(e_i),f_j\rangle}\\ &\underset{10.300}{=} \overline{\left\langle\sum_{k=1}^{m}M_{k,i}\cdot f_k,f_j\right\rangle}\\ &= \overline{\sum_{k=1}^{m}M_{k,i}\cdot\langle k_k,f_j\rangle}\\ &= \overline{\sum_{k=1}^{m}M_{k,i}\cdot\delta_{k,j}}\\ &= \overline{M_{j,i}}\end{aligned}$$

proving that $\forall(i,j)\in\{1,\ldots,n\}\times\{1,\ldots,m\}$ we have $\mathcal{M}(L^*,\{f_k\}_{k\in\{1,\ldots,m\}},\{e_k\}_{k\in\{1,\ldots,n\}})_{i,j}=\overline{\mathcal{M}(L,\{e_k\}_{k\in\{1,\ldots,n\}},\{f_k\}_{k\in\{1,\ldots,m\}})_{j,i}}$ □

Corollary 12.130. *Let $\langle X,\langle\rangle\rangle$ be a finite dimensional real inner product space, $L\in\mathrm{Hom}(X)$ then $\det(L)=\det(L^*)$*

Proof. Let $\{e_i\}_{i\in\{1,\ldots,n\}}$ be a basis in X then $\det(L)\underset{10.319}{=}\det(\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}}))\underset{10.318}{=}\det(\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}})^T)=\det(\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}})^H)\underset{12.129}{=}\det(L^*)$ □

12.5 Continuity

Definition 12.131. *Let $\langle X, \mathcal{T}_X \rangle$ and $\langle Y, \mathcal{T}_Y \rangle$ be topological spaces and let f: $X \rightarrow Y$ be a function between between X and Y then f is continuous at x iff $\forall V \in \mathcal{T}_Y \vdash f(x) \in V$ we have $\exists U \in \mathcal{T}_X \vdash x \in U \wedge f(U) \subseteq V$*

Theorem 12.132. *Let $\langle X, \mathcal{T}_X \rangle$ and $\langle Y, \mathcal{T}_Y \rangle$ be topological spaces with topological basis $\mathcal{B}_X$ and $\mathcal{B}_Y$ then for a function $f: X \rightarrow Y$ we have the following equivalences:*

1. *f is continuous at x*
2. *$\forall V \in \mathcal{B}_Y \vdash f(x) \in V$ we have $\exists U \in \mathcal{T}_X \vdash x \in U \wedge f(U) \subseteq V$*
3. *$\forall V \in \mathcal{B}_Y \vdash f(x) \in V$ we have $\exists U \in \mathcal{B}_X \vdash x \in U \wedge f(U) \subseteq V$*

Proof.

1. **($1 \Rightarrow 2$)** This is trivial as $\mathcal{B}_Y \subseteq \mathcal{T}_Y$
2. **($2 \Rightarrow 3$)** Let $V \in \mathcal{B}_Y$ then by (2) there exists a $U' \in \mathcal{T}_X$ with $x \in U' \wedge f(U') \subseteq V$, by the definition of a basis $\exists U \in \mathcal{B}_X$ with $x \in U \subseteq U' \Rightarrow f(U) \subseteq f(U') \subseteq V$
3. **($3 \Rightarrow 1$)** Let $V \in \mathcal{T}_Y$ with $f(x) \in V$ then by the definition of a basis there exists a $W \in \mathcal{B}_Y$ with $f(x) \in W \subseteq V$ and thus there $\exists U \in \mathcal{B}_X \subseteq \mathcal{T}_X$ such that $x \in U \wedge f(U) \subseteq W \subseteq V$ proving continuity □

Theorem 12.133. *Let $\langle X, \mathcal{T}_X \rangle$ and $\langle Y, \mathcal{T}_Y \rangle$ be topological spaces and let $f: X \rightarrow Y$ be a continuous function between X and Y then if $x \in A \subseteq X$ we have that $f_{|A}: A \rightarrow Y$ is continuous at x (using the subspace topology of A)*

Proof. If $V \in \mathcal{T}_Y$ with $f_A(x) \in V$ then we have $f(x) = f_{|A}(x) \in V$ and by continuity at x there exists a $W \in \mathcal{T}_X$ with $x \in W \wedge f(W) \subseteq V \underset{x \in A}{\Rightarrow} x \in W \bigcap A \subseteq W \Rightarrow f(W \bigcap A) \subseteq f(W) \subseteq V$ □

Theorem 12.134. *Let $\langle X, \mathcal{T}_X \rangle$ and $\langle Y, \mathcal{T}_Y \rangle$ be topological spaces, $f: X \rightarrow Y$, $f(X) \subseteq A$ then $f: X \rightarrow Y$ is continous on U iff $f: X \rightarrow A$ is continuous (using the subspace topology of A)*

Proof. If $f: X \rightarrow Y$ is continuous then if $x \in U$ and V open in A with $f(x) \in V$ then there exists a $W \in \mathcal{T}_Y$ such that $V = A \bigcap W \Rightarrow f(x) \in W$ so by continuity of f: $X \rightarrow Y$ there exists a $U \in \mathcal{T}_X$ with $x \in U$ and $f(U) \subseteq W \Rightarrow f(U) \underset{f(U) \subseteq A}{=} f(U) \bigcap A \subseteq W \bigcap A = V$ proving that $f: X \rightarrow A$ is continuous in the subspace topology at x and as x is choosen arbitrary $f: X \rightarrow A$ is continuous. If $f: X \rightarrow A$ is continous then if $x \in X$ and $V \in \mathcal{T}_Y$ such that $f(x) \in V$ then as $f(x) \in A$ we have $f(x) \in V \bigcap A$ so by continuity of $f: X \rightarrow A$ we have that $\exists U \in \mathcal{T}_X$ with $x \in U$ and $f(U) \subseteq V \bigcap A \subseteq V$ proving that $f: X \rightarrow Y$ is continuous at x and as x is choosen arbitrary it is $f: X \rightarrow Y$ is continuous. □

Definition 12.135. *Let $\langle X, \mathcal{T}_X \rangle$ and $\langle Y, \mathcal{T}_Y \rangle$ be topological spaces and let f: $X \rightarrow Y$ then f is continuous if and only if $\forall x \in X \vDash f$ is continuous at x. The set of continuous functions between X and Y is noted as $\mathcal{C}(X, Y)$.*

Example 12.136. Let $\langle X, \mathcal{T}_X \rangle$ and $\langle Y, \mathcal{T}_Y \rangle$ be topological functions, $c \in Y$ and $C_{X,c}: X \rightarrow Y$ defined by $C_c(x) = c$ is continuous

Proof. $\forall x \in X$ we have that if $c = C_{X,c}(x) \in V \in \mathcal{T}_Y$ then if $y \in X \in \mathcal{T}_X$ we have $C_{X,c}(y) = c \in V \Rightarrow C_{X,c}(X) \subseteq V$ □

Example 12.137. Let $\langle X, \mathcal{T} \rangle$ be a topological space then $1_X: X \rightarrow X$ defined by $x \rightarrow x$ is continuous

Proof. $\forall x \in X$ we have if $x = 1_X(x) \in V \in \mathcal{T}$ then if $y \in V$ then $1_X(y) = y \in V \Rightarrow 1_X(V) = V \subseteq V$ □

Theorem 12.138. *Let $\langle X, \mathcal{T}_X \rangle$ and $\langle Y, \mathcal{T}_Y \rangle$ be topological spaces and let $f: X \rightarrow Y$ be a function between X and Y then the following are equivalent*

1. *f is continuous*
2. *$\forall V \in \mathcal{T}_Y$ we have $f^{-1}(V) \in \mathcal{T}_X$*
3. *$\forall A \subseteq X$ we have $f(\bar{A}) \subseteq \overline{f(A)}$*
4. *$\forall F \subseteq Y \vdash F$ is closed we have that $f^{-1}(F)$ is closed*

Proof.

1. **(1 ⇒ 2)** Let $V \in \mathcal{T}_Y$ then either $f^{-1}(V) = \emptyset \in \mathcal{T}_X$ or $\forall x \in f^{-1}(V)$ we have by continuity that $\exists U_x \in \mathcal{T}_X \vdash x \in U_x \wedge f(U_x) \subseteq V \underset{2.54}{\Rightarrow} U_x \subseteq f^{-1}(V) \Rightarrow f^{-1}(V) = \bigcup_{x \in U} U_x \in \mathcal{T}_X$
2. **(2 ⇒ 1)** If $x \in X$ then if $f(x) \in V \subseteq \mathcal{T}_Y$ then $x \in U = f^{-1}(V) \in \mathcal{T}_X$ and $f(U) \subseteq V$
3. **(1 ⇒ 3)** Let $y \in f(\bar{A})$ then $\exists x \in \bar{A}$ such that $y = f(x)$, if now $y \in V \in \mathcal{T}_Y$ then by continuity $\exists U \in \mathcal{T}(x)$ such that $x \in U$ and $f(U) \subseteq V$, as $x \in \bar{A}$ we have by 12.19 that $U \bigcap A \neq \emptyset \Rightarrow \emptyset \neq f(U \bigcap A) \subseteq_{2.58} f(U) \bigcap f(A) \subseteq V \bigcap f(A) \Rightarrow y \in \overline{f(A)}$
4. **(3 ⇒ 2 and thus ⇒1)** Let A be a set in Y then as $f(\bar{A}) \subseteq \overline{f(A)}$ we have by 2.54 that $\bar{A} \subseteq f^{-1}(\overline{f(A)})$. So if $B \subseteq Y$ is closed we have $\bar{B}$=B and if we take $A = f^{-1}(B)$ then $\overline{f^{-1}(B)} = \bar{A} \subseteq f^{-1}(\overline{f(A)}) = f^{-1}(\overline{f(f^{-1}(B))}) \subseteq f^{-1}(\bar{B}) = f^{-1}(B)$, as we have also trivially $f^{-1}(B) \subseteq \overline{f^{-1}(B)}$ it follows that $f^{-1}(B) = \overline{f^{-1}(B)}$ and thus $f^{-1}(B)$ is closed. Now if $V \in \mathcal{T}_Y$ then $Y \setminus V$ is closed and as $X \setminus f^{-1}(V) \underset{2.54}{=} f^{-1}(Y \setminus V)$ is closed we have that $f^{-1}(V)$ is open.
5. **(2 ⇒ 4)** If F is closed then we have that $Y \setminus F$ is open and by (2) we have that $f^{-1}(Y \setminus U)$ is open so that $X \setminus f^{-1}(Y \setminus F)$ is closed. Now $X \setminus f^{-1}(Y \setminus F) \underset{2.54}{=} f^{-1}(Y \setminus (Y \setminus F)) = f^{-1}(F)$ so that $f^{-1}(F)$ is closed.
6. **(4 ⇒ 2)** If U is open in Y then $Y \setminus U$ is closed and by (4) we have that $f^{-1}(Y \setminus U)$ is closed and thus $X \setminus f^{-1}(Y \setminus U)$ is open. Finally $X \setminus f^{-1}(Y \setminus U) \underset{2.54}{=} f^{-1}(Y \setminus (Y \setminus U)) = f^{-1}(U)$ proving that $f^{-1}(U)$ is open. □

We prove now that continuity is essential a local feature

Theorem 12.139. *Let $\langle X, \mathcal{T}_X \rangle$, $\langle Y, \mathcal{T}_Y \rangle$ be topological spaces and $f: X \rightarrow Y$ a function between $X \rightarrow Y$ then f is continuous iff $\forall x \in X$ there exists a $U \in \mathcal{T}_X$ with $x \in U$ such that $f_{|U}: U \rightarrow Y$ is continuous in the subspace topology of U*

Proof.

1. **(⇒)** This follows from 12.133.
2. **(⇐)** Let $x \in X$ then there exists a $U \in \mathcal{T}_X$ with $x \in U$ such that $f_{|U}|U \to Y$ is continuous. So if V is open in Y containing $f_{|U}(x) = f(x)$ then there exists a W' open in U such that $f(W') = f_{|U}(W') \subseteq V$, as $W' \underset{\text{definition of subspace topology}}{=} W \bigcap U$ is open in X (as W, U are open) this means that f is continuous at x and as $x \in X$ was chosen arbitrary we have that f is continuous. □

Definition 12.140. *Let $\langle X, \mathcal{T}_X \rangle$, $\langle Y, \mathcal{T}_Y \rangle$ be topological spaces then a function $f: X \to Y$ is open iff $\forall U \in \mathcal{T}_X$ we have that $f(U) \in \mathcal{T}_Y$ (every image of a open set is open)*

Theorem 12.141. *Let $\langle X, \mathcal{T}_X \rangle$, $\langle Y, \mathcal{T}_Y \rangle$ be topological spaces then for a open injective function $f: X \to Y$ we have that $f^{-1}|f(X) \to X$ is continuous (see 2.19)*

Proof. Let U be open in X then as f is open we have that $f(U)$ is open and $(f^{-1})^{-1}(U) \underset{2.54}{=} f(U)$ is open in X and as $f(U) \subseteq f(X)$ we have that $f(U) \bigcap f(X) = f(U)$ so that $f(U) = (f^{-1})^{-1}(U)$ is open in $f(X)$ proving continuity of $f^{-1}: f(X) \to X$ □

Theorem 12.142. *Let $\langle X, \mathcal{T}_X \rangle$, $\langle Y, \mathcal{T}_Y \rangle$, $\langle Z, \mathcal{T}_Z \rangle$ be topological spaces and $f: X \to Y$, $g: Y \to Z$ be open functions then $g \circ f: X \to Z$ is a open function.*

Proof. This is trivial for if $U \in \mathcal{T}_X$ then $f(U)$ is open and thus $g(f(U)) = (g \circ f)(U)$ is open. □

Theorem 12.143. *Let $\langle X, \mathcal{T}_X \rangle$, $\langle Y, \mathcal{T}_Y \rangle$, $\langle Z, \mathcal{T}_Z \rangle$ be topological spaces and $f: X \to Y$, $g: Y \to Z$ be continuous functions then $g \circ f: X \to Z$ is continuous*

Proof. And let W be open in Z then by continuity of g we have that $g^{-1}(W)$ is open and thus by continuity of f we have that $(g \circ f)^{-1}(W) = f^{-1}(g^{-1}(W))$ is open. □

Theorem 12.144. *Let $\{\langle X_i, \mathcal{T}_i \rangle\}_{i \in I}$ be a family of topological spaces and let $\langle \prod_{i \in I} X_i, \mathcal{T} \rangle$ be the product of the family equipped with the product topology (see 12.36) then $\forall i \in I$ we have that the projection map $\pi_i: \prod_{j \in I} X_j \to X_i$ defined by $\pi_i(x_1, \ldots, x_n) = x_i$ is open and continuous*

Proof.

I. **(continuity)** If $U \in \mathcal{T}_i$ then by the definition of the product topology $\pi_i^{-1}(U)$ is a element of the generating sub-basis, thus a element of the basis and finally a element of the topology.

II. **(open)** Let U be open in the product topology and let $t \in \pi_i(U)$ then $\exists x \in U$ such that $t = \pi_i(x) = x(i) = x_i$, as $x \in U$ open we use 12.37 to find a family $\{A_i^{(t)}\}_{i \in I}$ and a finite $A \subseteq I$ such that $\forall j \in I$ we have $A_j^{(t)} = \begin{cases} A_j^{(t)} \text{ is open if } j \in A \\ A_j^{(t)} = X_j \text{ if } j \in I \setminus A \Rightarrow A_j^{(t)} \text{ is open} \end{cases}$ so that $x \in \prod_{i \in I} A_i^{(t)} \subseteq U$ so that $t = \pi_i(x) \subseteq \pi_i\left(\prod_{i \in I} A_i^{(t)}\right) \subseteq A_i^{(t)} \subseteq \pi_i(U)$ where A_i is open. Using 12.5 we have then that $\pi_i(U)$ is open. □

Theorem 12.145. *Let $\langle X, \mathcal{T}_X\rangle$ be a topological space, $\{X_i, \mathcal{T}_i\}_{i\in I}$ be a family of topological spaces and $\langle \prod_{i\in I} X_i, \mathcal{T}\rangle$ where $\mathcal{T}$ is the product topology. A function f: $X \rightarrow \prod_{i\in I} X_i$ is continuous iff $\forall i \in I$ we have that $f_i = \pi_i \circ f$ is continuous*

Proof.

1. **(⇒)** As f is continuous and $\forall i \in I$ we have π_i is continuous (see 12.144) we have by 12.143 that $f_i = \pi_i \circ f$ is continuous.

2. **(⇐)** Let $x \in X$ and $f(x) \in V$ which is open in the product topology (see 12.36), then by the definition of the product topology there exists a finite family $\{U_i^{(x)}\}_{i\in J}$ of open sets such that $f(x) \in \bigcap_{i\in J} \pi_i^{-1}(U_i^{(x)}) \subseteq V \Rightarrow x \in f^{-1}(\bigcap_{i\in J} \pi_i^{-1}(U_i^{(x)})) \subseteq f^{-1}(V) \underset{2.58}{\Rightarrow} x \in \bigcap_{i\in J} f^{-1}(\pi_i^{-1}(U_i^{(x)})) = \bigcap_{i\in J} (\pi_i \circ f)^{-1}(U_i^{(x)}) \subseteq f^{-1}(V)$. Now as $\pi_i \circ f$ is continuous we have that $W_x = \bigcap_{i\in J} (\pi_i \circ f)^{-1}(U_i^{(x)})$ is open and $x \in W_x \subseteq f^{-1}(V)$. Using 12.5 we have then that $f^{-1}(V)$ is open. □

Theorem 12.146. *Let $\langle X_1, \mathcal{T}_1\rangle$, $\langle X_2, \mathcal{T}_2\rangle$, $\langle Y, \mathcal{T}_Y\rangle$ be topological spaces, $\langle X_1 \times X_2, \mathcal{T}\rangle$ where $\mathcal{T}$ is the product topology (= the box topology by 12.39). If now $f: X_1 \times X_2 \rightarrow Y$ is a continuous function then the following functions are also continuous:*

1. *$\forall x \in X_1$ $f_1(x): X_2 \rightarrow Y$ defined by $t \rightarrow f_1(x)(t) = f(x, t)$*
2. *$\forall x \in X_2$ $f_2(x): X_1 \rightarrow Y$ defined by $t \rightarrow f_2(x)(t) = f(t, x)$*

Proof.

1. Let $V \subseteq Y$ be a open set then by continuity of f we have that $f^{-1}(V)$ is open in $X_1 \times X_2$. Take now $y \in f_1(x)^{-1}(V)$ then $f(x, y) = f_1(x)(y) \in V \Rightarrow (x, y) \in f^{-1}(V) \underset{f^{-1}(V) \text{ is open product topology is box topology}}{\Rightarrow} \exists U \in \mathcal{T}_1 \wedge \exists W \in \mathcal{T}_2$ such that $(x, y) \in U \times W \subseteq f^{-1}(V) \Rightarrow x \in U \wedge y \in W$. Now if $t \in W \Rightarrow (x, t) \in U \times W \subseteq f^{-1}(V) \Rightarrow f_1(x)(t) = f(x, t) \in V \Rightarrow t \in f_1(x)^{-1}(V) \Rightarrow W \subseteq f_1(x)^{-1}(V)$. So we have that $y \in W \subseteq f_1(x)^{-1}(V)$ which by 12.5 proves that $f_1(x)^{-1}(V)$ is open.

2. Let $V \subseteq Y$ be a open set then by continuity of f we have that $f^{-1}(V)$ is open in $X_1 \times X_2$. Take now $y \in f_2(x)^{-1}(V)$ then $f(y, x) = f_2(x)(y) \in V \Rightarrow (y, x) \in f^{-1}(V) \underset{f^{-1}(V) \text{ is open product topology is box topology}}{\Rightarrow} \exists U \in \mathcal{T}_1 \wedge \exists W \in \mathcal{T}_2$ such that $(y, x) \in U \times W \subseteq f^{-1}(V) \Rightarrow y \in U \wedge x \in W$. Now if $t \in U \Rightarrow (t, x) \in U \times W \subseteq f^{-1}(V) \Rightarrow f_2(x)(t) = f(t, x) \in V \Rightarrow t \in f_2(x)^{-1}(V) \Rightarrow U \subseteq f_1(x)^{-1}(V)$. So we have that $y \in U \subseteq f_2(x)^{-1}(V)$ which by 12.5 proves that $f_2(x)^{-1}(V)$ is open. □

We now prove the condition to have a continuous function in metric spaces. From now on we assume that for the metric spaces we use the metric topology.

Theorem 12.147. *Let $\langle X, d_X\rangle$, $\langle Y, d_Y\rangle$ be metric spaces and $f: X \rightarrow Y$ a function then f is continuous in the metric topology at $x \in X$ iff $\forall \varepsilon \in \mathbb{R}_+$ we have that $\exists \delta \in \mathbb{R}_+$ such that $\forall y \in X \vdash d_X(x, y) < \delta$ we have $d_Y(f(x), f(y)) < \varepsilon$*

Proof.

1. **(⇒)** If f is continuous at x then if $\varepsilon\in\mathbb{R}_+$ we have that $f(x)\in B_{d_Y}(f(x),\varepsilon)$ which is open so there exists a U open in X such that $x\in U$ and $f(U)\subseteq B_{d_Y}(f(x),\varepsilon)$. As U is open in the metric topology there exists a $\delta\in\mathbb{R}_+$ such that $x\in B_{d_X}(x,\delta)\subseteq U\Rightarrow f(B_{d_X}(x,\delta))\subseteq f(U)\subseteq B_{d_Y}()$
2. **(⇐)** Let $x\in X$ then to prove continuity we can use 12.132 so if $\varepsilon\in\mathbb{R}_+$ and $f(x)\in B_{d_Y}(f(x),\varepsilon)$. Now by the hypothesis there exists a $\delta\in\mathbb{R}_+$ such that if $d_X(x,y)<\delta\Rightarrow d_Y(f(x),f(y))<\varepsilon$ or if $y\in B_{d_X}(x,\delta)\Rightarrow d_X(x,y)<\delta\Rightarrow d_Y(f(x),f(y))\Rightarrow f(y)\in B_{d_Y}(f(x),\varepsilon)\Rightarrow f(B_{d_X}(x,\delta))\subseteq B_{d_Y}(y,\delta)$ □

Corollary 12.148. *Let $\langle X,d_X\rangle$, $\langle Y,d_Y\rangle$ be metric space, $A\subseteq X$ then $f:A\to Y$ is continuous (using the subspace topology on A based on the metric topology of X) at $x\in A$ iff $\forall\varepsilon\in\mathbb{R}_+$ there exists a $\delta\in\mathbb{R}_+$ such that $\forall y\in A$ with $d_X(x,y)<\delta$ we have $d_Y(f(x),f(y))$*

Proof. Using 12.57 we have that the sub space topology is generated by the restricted metric $d_{X|A\times A}:A\times A\to\mathbb{R}$ defined by $d_{X_{|A\times A}}(x,y)=d_A(x,y)$. Using the previous theorem we have then that $f:A\to Y$ is continuous iff $\forall\varepsilon\in\mathbb{R}_+$ there exists a $\delta\in\mathbb{R}_+$ such that $\forall y\in A$ with $d_{X_{|A\times A}}(x,y)<\varepsilon$ we have $d_Y(f(x),f(y))<\varepsilon$ which as $d_{X_{|A\times A}}(x,y)=d_X(x,y)$ is equivalent with $\forall\varepsilon\in\mathbb{R}_+$ there exists a $\delta\in\mathbb{R}_+$ such that $\forall y\in A$ with $d_X(x,y)<\varepsilon$ we have $d_Y(f(x),f(y))<\varepsilon$ □

Theorem 12.149. *Let $\langle X,d_X\rangle$ and $\langle Y,d_Y\rangle$ be metric spaces and $f:X\to Y$ a function then f is open if and only if $\forall x\in X,\forall\delta>0$ we have that there exists a $\delta'>0$ such that $f(x)\in B_{d_Y}(f(x),\delta')\subseteq f(B_{d_X}(x,\delta))$*

Proof.

1. **(⇒)** Because f is open we have that if $x\in B_{d_X}(x,\delta)$ then $f(B_{d_X}(x,\delta))$ is open and thus as $f(x)\in f(B_{d_X}(x,\delta))$ there exists a δ' such that $f(x)\in B_{d_Y}(f(x),\delta')\subseteq f(B_{d_X}(x,\delta))$
2. **(⇐)** Let $U\subset X$, U open and let $y\in f(U)$ then there exists a $x\in U$ such that $f(x)=y$ then by the hypothesis there exists a δ_y with $y=f(x)\in B_{d_Y}(f(x),\delta_y)=B_{d_Y}(y,\delta_y)\subseteq f(B_{d_X}(x,\delta))\subseteq f(U)\Rightarrow f(U)=\bigcup_{y\in f(U)}B_{d_Y}(y,\delta_y 0$ and thus $f(U)$ is open □

Definition 12.150. *Let $\langle X,d_X\rangle$, $\langle Y,d_Y\rangle$ be metric spaces then a function f: $X\to Y$ is **uniform continuous** in $K\subseteq X$ iff $\forall\varepsilon\in\mathbb{R}_+$ we have that $\exists\delta\in\mathbb{R}_+$ such that $\forall x,y\in K$ with $d_X(x,y)<\varepsilon$ we have that $d_Y(f(x),f(y))<\varepsilon$*

Theorem 12.151. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces then a function f: $X\to Y$ is continuous at x in the norm topologies iff $\forall\varepsilon\in\mathbb{R}_+$ there exists a $\delta\in\mathbb{R}_+$ such that if $\|x-y\|_X<\delta$ then $\|f(x)-f(y)\|_Y<\varepsilon$*

Proof. This follows from 12.147 and the fact that $d_X(x,y)=\|x-y\|_X$ and $d_Y(f(x),f(y))=\|f(x)-f(y)\|$ □

We can extend the above theorem to a subspace of a normed space.

Theorem 12.152. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces, $A \subseteq X$ $A \subseteq X$ then $f: A \rightarrow Y$ is continuous (using the subspace topology on A based on the normed topology of X) at $x \in A$ iff $\forall \varepsilon \in \mathbb{R}_+$ there exists a $\delta \in \mathbb{R}_+$ such that $\forall y \in A$ with $\|x - y\|_X < \delta$ we have $\|f(x) - f(y)\| < \varepsilon$*

Proof. This follows from 12.148 and the fact that $d_X(x, y) = \|x - y\|_X$ and $d_Y(f(x), f(y)) = \|f(x) - f(y)\|$ □

Definition 12.153. *Let $\langle X, \|\|_X\rangle$ and $\langle Y, \|\|_Y\rangle$ be normed spaces then a function $f: X \rightarrow Y$ is **Lipschitz continuous** if $\forall x, y \in X$ we have $\|f(x) - f(y)\|_Y \leqslant \|x - y\|_X$*

Theorem 12.154. *Let $\langle X, \|\|_X\rangle$ and $\langle Y, \|\|_Y\rangle$ be normed spaces then a Lipschitz continuous function $f: X \rightarrow Y$ is uniform continuous and thus continuous.*

Proof. Let $\varepsilon > 0$ take $d_{\|\|_X}(x, y) < \varepsilon$, then we have $d_{\|\|_Y}(f(x) - f(y)) = \|f(x) - f(y)\|_Y \leqslant \|x - y\|_X = d_{\|\|_X}(x, y) < \varepsilon \Rightarrow d_{\|\|_Y}(f(x, f(y)))$ □

Definition 12.155. *Let $\langle X, \|\|\rangle$ be a normed space, $\emptyset \neq S \subseteq X$ a non-empty set then the **set distance function** $\delta_S: X \rightarrow \mathbb{R}$ is defined by $x \rightarrow \delta_S(x) = \inf(\{\|x - y\| | y \in S\})$. Note that if $x \in S$ then $\delta_S(x) = 0$.*

Theorem 12.156. *Let $\langle X, \|\|\rangle$ be a normed space, $\emptyset \neq S \subseteq X$ a non-empty set then $\delta_S: X \rightarrow \mathbb{R}$ is Lispchitz continuous and thus continuous.*

Proof. Let $x, y \in X$, $\varepsilon > 0$ then there exists a $y_\varepsilon \in S$ such that $\inf(\{\|x - y\| | y \in S\}) \leqslant \|x - y_\varepsilon\| < \inf(\{\|x - y\| | y \in S\}) + \varepsilon \Rightarrow \delta_S(x) \leqslant \|x - y_\varepsilon\| < \delta_S(x) + \varepsilon$. Then $\delta_S(y) - \delta_S(x) \leqslant \|y - y_\varepsilon\| - \|x - y_\varepsilon\| + \varepsilon \leqslant \|y - x\| + \|x - y_\varepsilon\| - \|x - y_\varepsilon\| + \varepsilon = \|y - x\| + \varepsilon \Rightarrow \delta_S(y) - \delta_S(x) \leqslant \|y - x\| + \varepsilon$. If now $\|y - x\| < \delta_S(y) - \delta_S(x)$ take then $\varepsilon = \frac{\delta_S(y) - \delta_S(x) - \|y - x\|}{2} > 0$ then we have $\delta_S(y) - \delta_S(x) \leqslant \|y - x\| + \varepsilon \leqslant \|y - x\| + \frac{\delta_S(y) - \delta_S(x) - \|y - x\|}{2} < \|y - x\| + \delta_S(y) - \delta_S(x) - \|y - x\| = \delta_S(y) - \delta_S(x)$ giving the contradiction $\delta_S(y) - \delta_S(x) < \delta_S(y) = \delta_S(x)$, so we must have

$$\delta_S(y) - \delta_S(x) \leqslant \|y - x\|$$

Exchange x, y gives then also

$$\delta_S(x) - \delta_S(y) \leqslant \|x - y\| = \|y - x\|$$

proving that $|\delta_S(y) - \delta_S(x)| \leqslant \|y - x\|$ □

Theorem 12.157. *Let $\langle X, \|\|\rangle$ be a normed space then $+: X \times X \rightarrow X$ defined by $(x, y) \rightarrow x + y$, $\cdot: X \times \mathbb{R}(\mathbb{C}) \rightarrow X$ defined by $(x, \alpha) \rightarrow \alpha \cdot x$ and $\|\|: X \rightarrow \mathbb{R}$ are continuous in the norm topology and product topologies of $X \times X$ and $X \times \mathbb{R}(\mathbb{C})$*

Proof.

1. **(+)** Let $x + y \in U$ (open in the norm topology) then $\exists \varepsilon \in \mathbb{R}_+$ such that $x + y \in B_{d_{\|\|}}(x + y, \varepsilon) \subseteq U$. Take now $z' = (x', y') \in B_{d_{\|\|}}\left(x, \frac{\varepsilon}{2}\right) \times B_{d_{\|\|}}\left(y, \frac{\varepsilon}{2}\right)$ (which is open in $X \times X$) then $\|x + y - (x' + y')\| = \|(x - x') + (y - y')\| \leqslant \|x - x'\| + \|y - y'\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \Rightarrow x' + y' \in B_{d_{\|\|}}(x + y, \varepsilon) \subseteq U \Rightarrow +\left(B_{d_{\|\|}}\left(x, \frac{\varepsilon}{2}\right) \times B_{d_{\|\|}}\left(y, \frac{\varepsilon}{2}\right)\right) \subseteq U$ proving that $+$ is continue at (x, y). As $(x, y) \in X \times X$ has been chosen arbitrary we have that $+$ is continuous.

2. **(.)** Let $\lambda \cdot x \in U$ (open in X) then there exists a $\varepsilon \in \mathbb{R}_+$ such that $\lambda \cdot x \in B_{d_{\|\|}}(\lambda \cdot x, \varepsilon) \subseteq U$. Let know $z' = (x', \lambda') \in B_{d_{\|\|}}(x, \delta_1) \times B_{d_{||}}(\lambda, \delta_2)$ which is open in $X \times \mathbb{R}(\mathbb{C})$ then $\|\lambda \cdot x - \lambda' \cdot x'\| = \|\lambda \cdot x - \lambda' \cdot x + \lambda' \cdot x - \lambda' \cdot x'\| \leqslant \|\lambda \cdot x - \lambda' \cdot x\| + \|\lambda' \cdot x - \lambda' \cdot x'\| = |\lambda - \lambda'| \cdot \|x\| + |\lambda'| \cdot \|x - x'\| < \delta_2 \cdot \|x\| + |\lambda'| \cdot \delta_1 = \delta_2 \cdot \|x\| + \delta_1 \cdot |\lambda' - \lambda + \lambda| \leqslant \delta_2 \cdot \|x\| + \delta_1 \cdot (|\lambda' - \lambda| + |\lambda|) < \delta_2 \cdot \|x\| + \delta_1 \cdot (\delta_2 + |\lambda|)$ giving
$$\|\lambda \cdot x - \lambda' \cdot x'\| < \delta_2 \cdot \|x\| + \delta_1 \cdot (\delta_2 + |\lambda|) \tag{12.16}$$
consider now the following cases
   a. **($\|\boldsymbol{x}\| = \mathbf{0}$)** take then $\delta_2 = 1$ and $\delta_1 = \frac{\varepsilon}{1 + |\lambda|}$ then $\delta_2 \cdot \|x\| + \delta_1 \cdot (\delta_2 + |\lambda|) = \frac{\varepsilon}{1 + |\lambda|} \cdot (1 + |\lambda|) = \varepsilon \underset{12.16}{\Rightarrow} \|\lambda \cdot x - \lambda' \cdot x'\| < \varepsilon$
   b. **($\|\boldsymbol{x}\| \neq \mathbf{0}$)** take then $\delta_1 = \frac{\varepsilon}{2 \cdot (1 + |\lambda|)} \wedge \delta_2 = \min\left(1, \frac{\varepsilon}{2 \cdot \|x\|}\right)$ then $\delta_2 \cdot \|x\| + \delta_1 \cdot (\delta_2 + |\lambda|) \leqslant \frac{\varepsilon}{2 \cdot \|x\|} \cdot \|x\| + \frac{\varepsilon}{2 \cdot (1 + |\lambda|)} \cdot (1 + |\lambda|) = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \underset{12.16}{\Rightarrow} \|\lambda \cdot x - \lambda' \cdot x'\| < \varepsilon$

   So with the correct choice of δ_1, δ_2 we have $\|\lambda \cdot x - \lambda' \cdot x'\| < \varepsilon \Rightarrow \lambda' \cdot x' \in B_{d_{\|\|}}(\lambda \cdot x, \varepsilon) \subseteq U$ if $(x', \lambda') \in B_{d_{\|\|}}(x, \delta_1) \cdot B_{d_{||}}(\lambda, \delta_2)$ or $\cdot(B_{d_{\|\|}}(x, \delta_1) \cdot B_{d_{||}}(\lambda, \delta_2)) \subseteq U$ proving that $\cdot$ is continuous at (x, λ). As (x, λ) was chosen arbitrary we have that $\cdot$ is continuous.

3. **($\|\|$)** Let $x \in X$ and let $\varepsilon \in \mathbb{R}_+$ then if $\|x - y\| < \varepsilon \Rightarrow |\|x\| - \|y\|| \leqslant \|x - y\| < \varepsilon$ proving by 12.151 that $\|\|$ is continue at x. As x is chosen arbitrary we have that $\|\|$ is continuous. □

Corollary 12.158. *Let $\langle X, \|\|\rangle$ be a normed space then we have*

1. *$\forall x \in X$ the function $\tau_x: X \rightarrow X$ defined by $\tau_x(y) = x + y$ is continuous*
2. *$\forall \alpha \in \mathbb{R}(\mathbb{C})$ the function $\mu_\alpha: X \rightarrow X$ defined by $\mu_\alpha(y) = \alpha \cdot y$ is continuous*
3. *$\forall x \in X$ the function $\nu_x: \mathbb{R}(\mathbb{C}) \rightarrow X$ defined by $\nu_x(\alpha) = \alpha \cdot x$ is continuous*

Proof. This follows from the continuity of $+: X \times X \rightarrow X$ and $.: X \times \mathbb{R}(\mathbb{C}) \rightarrow X$ aand 12.146. □

Theorem 12.159. *When $\langle \mathbb{R}^n, \|\|\rangle$ is equipped with product topology generated by the maximum norm $\|\|$ (see 12.79) and $\|\|^*$ is another norm on $\mathbb{R}^n$ then the function $\|\|^*: \mathbb{R}^n \rightarrow \mathbb{R}$ defined by $x \rightarrow \|x\|^*$ is continuous.*

Proof. Let $\{\mathcal{E}_i\}_{i \in \{1, \ldots, n\}}$ be the canonical basis on $\mathbb{R}^n$ (see 10.161) then $\forall x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ we have $x = \sum_{i=1}^n x_i \cdot \mathcal{E}_i \Rightarrow \|x\|^* = \|\sum_{i=1}^n x_i \cdot \mathcal{E}_i\|^* \leqslant_{12.70} \sum_{i=1}^n \|x_i \cdot \mathcal{E}_i\|^* = \sum_{i=1}^n |x_i| \cdot \|\mathcal{E}_i\|^* \leqslant \max(\{|x_i| | i \in \{1, \ldots, n\}\}) \cdot \sum_{i=1}^n \|\mathcal{E}_i\|^* = \|x\| \cdot \sum_{i=1}^n \|\mathcal{E}_i\|^* = A \cdot \|x\|$ where $A = \sum_{i=1}^n \|\mathcal{E}_i\|^* \geqslant 0$. So we have
$$\|x\|^* \leqslant A \cdot \|x\| \text{ where } A \geqslant 0 \tag{12.17}$$
If now $x \in \mathbb{R}^n$ and $\varepsilon \in \mathbb{R}_+$ take then $\delta = \frac{\varepsilon}{A + 1}$ then if $\|x - y\| < \delta \Rightarrow \|x - y\|^* \leqslant A \cdot \|x - y\| < A \cdot \delta = A \cdot \frac{\varepsilon}{A + 1} \leqslant \varepsilon$ (as $0 \leqslant A$) proving that $\|\|^*$ is continuous at x. As x was chosen arbitrary we have proved that $\|\|^*$ is continuous. □

Definition 12.160. *Let $\langle X, \mathcal{T}_X\rangle$, $\langle Y, \mathcal{T}_Y\rangle$ be topological spaces then a function $f: X \rightarrow Y$ is a **homeomorphism** iff f is a bijection and f, f^{-1} are continuous. $\langle X, \mathcal{T}_X\rangle$ and $\langle Y, \mathcal{T}_Y\rangle$ are then said to be **homeomorphic**.*

Note 12.161. Using 12.141 it is easy to see that $f: X \rightarrow Y$ is a homeomorphism if it is a bijection and if it is open and continuous.

Theorem 12.162. *Let $\langle X, \mathcal{T}_X\rangle$, $\langle Y, \mathcal{T}_Y\rangle$ be topological spaces, $f: X \rightarrow Y$ a homeomorphism and $\emptyset \neq A \subseteq X$ then $f_{|A}: A \rightarrow f(A)$ is a homeomorphism (using the subspace topologies of $\mathcal{T}_X$ and $\mathcal{T}_Y$ on A and $f(A)$.*

Proof.

1. **(bijectivity)** First using 2.49 and the fact that f is bijective and thus injective we have that $f_{|A}: A \rightarrow f(A)$ is a bijection.
2. **(continuity)** Let $x \in A$ then if V is a open set in $f(A)$ containing $f_{|A}(x) = f(x)$ there exists a $W \in \mathcal{T}_Y$ such that $V = f(A) \bigcap W$. As f is continuous at x there exists a $U \in \mathcal{T}_X$ with $x \in U$ such that $f(U) \subseteq W$. So we have $x \in A \bigcap U$ (which is open by definition in A) and $f_{|A}(A \bigcap U) \underset{A \bigcap U \subseteq A}{=} f(A \bigcap U) \underset{f \text{ is injective and } 2.58}{=} f(A) \bigcap f(U) \subseteq f(A) \bigcap W = V$ proving that $f_{|A}$ is continous at x. As $x \in A$ was choosen arbitrary we conclude that $f_{|A}$ is continuous.
3. **(open)** Let U open in A then there exists a $W \in \mathcal{T}_X$ such that $U = A \bigcap W$. We have then that $f_{|A}(U) \underset{U \subseteq A}{=} f(U) = f(A \bigcap W) \underset{f \text{ is injective and } 2.58}{=} f(A) \bigcap f(W)$ as f is open we have that $f(W) \in \mathcal{T}_Y \Rightarrow f_{|A}(U) = f(A) \bigcap W$ is open in the subspace topology of $f(A)$. □

Theorem 12.163. *Let $\langle X, \mathcal{T}_X\rangle$, $\langle Y, \mathcal{T}_Y\rangle$ be topological spaces then if $f: X \rightarrow Y$ is a homeomorphism we have that $f^{-1}: Y \rightarrow X$ is a homeomorphism*

Proof. As f is a bijection we have by 2.38 that f^{-1} is a bijection. By definition we have that f^{-1} is continuous and by 2.41 we have that $(f^{-1})^{-1} = f$ is continuous proving that f^{-1} is a homeomorphism. □

Theorem 12.164. *Let $\langle X, \mathcal{T}_X\rangle$, $\langle Y, \mathcal{T}_Y\rangle$ and $\langle Z, \mathcal{T}_Z\rangle$ be topological spaces and $f: X - Y$, $g: Y \rightarrow Z$ be homeomorphism's's then $g \circ f: X \rightarrow Z$ is a homeomorphism.*

Proof. Using 2.46 we have that $g \circ f$ is a bijection, also using 12.161 we have that g, f are continuous and open. Using 12.143,12.142 we have then that $g \circ f$ is open and continuous, which by 12.161 means that $g \circ f$ is a homeomorphism. □

Theorem 12.165. *Let $\langle X, \mathcal{T}\rangle$ be a topological space, Y a set and $f: X \rightarrow Y$ a bijection. Define then $\mathcal{T}_{f,\mathcal{T}} = \{f(U) | U \in \mathcal{T}\}$ forms a topology on Y so that f is a homeomorphism. Further if $g: X \rightarrow Y$ is such that $f = g \circ h$ where $h: X \rightarrow X$ is a homeomorphism then g is a homeomorphism and $\mathcal{T}_{f,\mathcal{T}} = \mathcal{T}_{g,\mathcal{T}}$*

Proof.

1. **($\mathcal{T}_{f,\mathcal{T}}$ is a topology)** We have to full fill all the axioms of a topology
   a. $\emptyset = f(\emptyset) \in \mathcal{T}_{f,\mathcal{T}}$

b. $Y \underset{f \text{ is bijection thus surjective}}{=} f(X) \in \mathcal{T}_{f,\mathcal{T}}$

c. If $\{V_i\}_{i \in I}$ is a family of sets in $\mathcal{T}_{f,\mathcal{T}}$ then $\forall i \in I$ there exists a $U_i \in \mathcal{T}$ such that $V_i = f(U_i)$ so that $\bigcup_{i \in I} V_i = \bigcup_{i \in I} f(U_i) \underset{2.58}{=} f(\bigcup_{i \in I} U_i) \in \mathcal{T}_{f,\mathcal{T}}$

d. Let $V_1, V_2 \in \mathcal{T}_{f,\mathcal{T}}$ then $\exists U_1, U_2 \in \mathcal{T}$ such that $V_1 = f(U_1)$, $V_2 = f(U_2)$ then $V_1 \bigcap V_2 = f(U_1) \bigcap f(U_2) \underset{2.58}{=} f(U_1 \bigcap U_2) \in \mathcal{T}_{f,X}$

2. **(f is a homeomorphism)** If $U \in \mathcal{T}$ then $f(U) \in \mathcal{T}_{f,\mathcal{T}}$ so that f is open. If $V \in \mathcal{T}_{f,\mathcal{T}}$ then $\exists U \in \mathcal{T}$ such that $V = f(U) \Rightarrow f^{-1}(V) = f^{-1}(f(U)) \underset{2.52}{=} U \in \mathcal{T}$ proving that f is continuous. As f is a bijection we have that f is a homeomorphism.

3. **($\mathcal{T}_{f,X} = \mathcal{T}_{g,X}$)** First as h is a homeomorphism and thus a bijection we have that $h^{-1}: X \rightarrow X$ is a homeomorphism (see 12.163) so that $f \circ h^{-1} = (g \circ h) \circ h^{-1} = g \circ (h \circ h^{-1}) = g \circ 1_X = g$ so that g is a homeomorphism (see 12.164) and $\mathcal{T}_{g,\mathcal{T}}$ is defined. Next if $U \in \mathcal{T}_{f,\mathcal{T}}$ then $\exists V \in \mathcal{T}$ such that $U = f(V) = f(h^{-1}(h(V))) = (f \circ h^{-1})(h(V)) = g(h(V)) \in \mathcal{T}_{g,\mathcal{T}}$ as h is a homeomorphism so that $g(V) \in \mathcal{T}$. Finally if $U \in \mathcal{T}_{g,\mathcal{T}}$ then $\exists V \in \mathcal{T}$ such that $U = g(V) = g(h(h^{-1}(V))) = (g \circ h)(h^{-1}(V)) = f(h^{-1}(V)) \in \mathcal{T}_{g,\mathcal{T}}$ as $h^{-1}(V) \in \mathcal{T}$ because h is a homeomorphism. □

Theorem 12.166. *Let X be a set, $\langle Y, T\rangle$ a topological space and $f: X \rightarrow Y$ a function then $\mathcal{T}_{f,\mathcal{T}}^{-1} = \{f^{-1}(U) | U \in \mathcal{T}\}$ is a topology on X called the inverse induced topology. Furthermore f is a continuous function between X and Y if X is equipped with the inverse induced topology.*

Proof.

1. $\emptyset = f^{-1}(\emptyset) \in \mathcal{T}_{f,\mathcal{T}}^{-1}$
2. $X = f^{-1}(Y) \in \mathcal{T}_{f,\mathcal{T}}^{-1}$
3. Let $\{V_i\}_{i \in I}$ be a family of sets with $\forall i \in I$ we have $V_i \in \mathcal{T}_{f,\mathcal{T}}^{-1}$ so that $\exists U_i \in \mathcal{T}$ with $V_i = f^{-1}(U_i)$ then $\bigcup_{i \in I} V_i = \bigcup_{i \in I} f^{-1}(U_i) \underset{2.58}{=} f^{-1}(\bigcup_{i \in I} U_i) \in \mathcal{T}_{f,\mathcal{T}}^{-1}$ [as $\bigcup_{i \in I} U_i \in \mathcal{T}$].
4. If $V_1, V_2 \in \mathcal{T}_{f,\mathcal{T}}^{-1}$ then $\exists U_1, U_2 \in \mathcal{T}$ such that $f^{-1}(U_1) = V_1$, $f^{-1}(U_2) = V_2$. So $V_1 \bigcap V_2 = f^{-1}(U_1) \bigcap f^{-1}(U_2) \underset{2.58}{=} f^{-1}(U_1 \bigcap U_2) \in \mathcal{T}_{f,\mathcal{T}}^{-1}$ [as $U_1 \bigcap U_2 \in \mathcal{T}$]

As by definition we have if $U \in \mathcal{T}$ that $f^{-1}(U) \in \mathcal{T}_{f,\mathcal{T}}^{-1}$ so that f is continuous. □

Theorem 12.167. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed vector spaces over $\mathbb{R}(\mathbb{C})$ and let $f: X \rightarrow Y$ be a linear isometry then f is a homeomorphism.*

Proof. As f is a isometry it is a bijection. To prove continuity take $\varepsilon \in \mathbb{R}_+$ and $x \in X$ then if $\|x - x'\|_X < \varepsilon$ we have that $\varepsilon > \|x - x'\|_X = \|f(x - x')\|_Y = \|f(x) - f(x')\|_Y$ so that by 12.151 f is continuous. As f^{-1} is a isometry by 12.85 and by 10.181 is also linear we can apply the above reasoning to prove that f^{-1} is continuous. □

Example 12.168. Let $\langle \mathbb{R}^2, \|\|_e\rangle$ be equiped with the euclidean norm (see 12.107), $\langle \mathbb{C}, ||\rangle$ the complex space with the canonical norm then $\mathcal{C}: \mathbb{R}^2 \rightarrow \mathbb{C}$ is a homeomorphism

Proof. Using 10.180 we have that $\mathcal{C}$ is a isomorphism and by 12.109 a isometry and thus a linear isometry. Applying then the above theorem (see 12.167 we have that $\mathcal{C}$ is a homeomorphism. $\square$

Theorem 12.169. *Let $\langle X, \|\|\rangle$ be a finite n-dimensional normed space over $\mathbb{R}(\mathbb{C})$ then there exist a norm $\|\|_X$ on $\mathbb{R}^n(\mathbb{C}^n)$ such that the function $\varphi: X \to \mathbb{R}^n(\mathbb{C}^n)$ defined by $x = \sum_{i \in \{1,\ldots,n\}} x_i \cdot e_i \to \varphi(x) = (x_1, \ldots, x_n)$ (where $\{e_i\}_{i \in I\{1,\ldots,n\}}$ is a basis in X) is a isometry, isomorphism and by the previous theorem they also a homeomorphism.*

Proof. First we prove that φ is a bijection

1. **(injectivity)** If $\varphi(x) = \varphi(x')$ where $x = \sum_{i \in \{1,\ldots,n\}} x_i \cdot e_i$, $x' = \sum_{i \in \{1,\ldots,n\}} x_i' \cdot e_i$ are the unique expansions in the basis $\{e_i\}_{i \in I}$ then $(x_1, \ldots, x_n) = (x_1', \ldots, x_n') \Rightarrow \forall i \in I \vDash x_i = x_i'$ and thus $x = x'$.
2. **(surjectivity)** If $(x_1, \ldots, x_n) \in \mathbb{R}^n(\mathbb{C}^n)$ then $\varphi\left(\sum_{i \in \{1,\ldots,n\}} x_i \cdot e_i\right) = (x_1, \ldots, x_n)$

Next to prove linearity if $x = \sum_{i \in \{1,\ldots,n\}} x_i \cdot e_i$, $y = \sum_{i \in \{1,\ldots,n\}} y_i \cdot e_i$ and $\alpha, \beta \in \mathbb{R}(\mathbb{C})$ then $\alpha \cdot x + \beta \cdot y = \alpha \cdot \sum_{i \in \{1,\ldots,n\}} x_i \cdot e_i + \beta \cdot \sum_{i \in \{1\ldots,n\}} y_i \cdot e_i = \sum_{i \in \{1,\ldots,n\}} (\alpha \cdot x_i + \beta \cdot y_i) \Rightarrow \varphi(\alpha \cdot x + \beta \cdot y) = \alpha \cdot \varphi(x) + \beta \cdot \varphi(y)$.

Define now the norm $\|\|_X$ on $\mathbb{R}^n(\mathbb{C}^n)$ by $\|(x_1, \ldots, x_n)\|_X = \left\|\sum_{i \in \{1,\ldots,n\}} x_i \cdot e_i\right\|$ then $\|\|$ is indeed a norm:

1. $\|(x_1, \ldots, x_n)\|_X = \left\|\sum_{i \in \{1,\ldots,n\}} x_i \cdot e_i\right\| \geqslant 0$
2. $\|\alpha \cdot (x_1, \ldots, x_n)\|_X = \|(\alpha \cdot x_1, \ldots, \alpha \cdot x_n)\|_X = \left\|\sum_{i \in \{1,\ldots,n\}} (\alpha \cdot x_i) \cdot e_i\right\|_X = \left\|\alpha \cdot \sum_{i \in \{1,\ldots,n\}} x_i \cdot e_i\right\| = |\alpha| \cdot \left\|\sum_{i \in \{1,\ldots,n\}} x_i \cdot e_i\right\| = |\alpha| \cdot \|x\|$
3. $\|(x_1, \ldots, x_n) + (y_1, \ldots, y_n)\|_X = \|(x_1 + y_1, \ldots, x_n + y_n)\|_X = \left\|\sum_{i \in \{1,\ldots,n\}} (x_i + y_i)\right\| = \left\|\sum_{i \in \{1,\ldots,n\}} x_i + \sum_{i \in \{1,\ldots,n\}} y_i\right\| \leqslant \left\|\sum_{i \in \{1,\ldots,n\}} x_i\right\| + \left\|\sum_{i \in \{1,\ldots,n\}} y_i\right\| = \|(x_1, \ldots, x_n)\|_X + \|(y_1, \ldots, y_n)\|_X$

If now $x = \sum_{i \in \{1,\ldots,n\}} x_i \cdot e_i$ then $\|\varphi(x)\|_X = \|(x_1, \ldots, x_n)\|_X = \left\|\sum_{i \in \{1,\ldots,n\}} x_i \cdot e_i\right\| = \|x\|$ proving that φ is a isometry. $\square$

Theorem 12.170. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces over $\mathbb{R}(\mathbb{C})$, $U \subseteq X$ a open set, $V_f, V_g \subseteq Y$ open sets, $f: U \to V_f$, $g: U \to V_g$ be continuous functions and $\alpha \in \mathbb{R}(\mathbb{C})$ then $f + g: U \to V_f + V_g$ and $\alpha \cdot f: U \to V_f$ are continuous.*

Proof. We use 12.151 to prove continuity at every point $x \in U$ and thus continuity. So take $\varepsilon \in \mathbb{R}_+$ and $x \in X$

1. By continuity of f, g there exists a $\delta_1, \delta_2 \in \mathbb{R}_+$ such that if $\|x - x'\|_X < \delta_1$ then $\|f(x) - f(x')\|_Y < \frac{\varepsilon}{2}$ and if $\|x - x'\|_X < \delta_2$ then $\|f(x) - f(x')\|_Y < \frac{\varepsilon}{2}$. Take now $\delta = \min(\delta_1, \delta_2)$ then if $\|x - x'\|_X < \delta \Rightarrow \|x - x'\|_X < \delta_1, \delta_2 \Rightarrow \|(f + g)(x) - (f + g)(x')\|_Y = \|f(x) - f(x') + g(x) - g(x')\|_Y \leqslant \|f(x) - f(x')\|_Y + \|g(x) - g(x')\|_Y = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$

2. Consider the following cases for $\alpha \in \mathbb{R}(\mathbb{C})$

   a. $(\boldsymbol{\alpha = 0})$ then if $\|x - x'\|_X < 1$ we have that $\|0 \cdot f(x) - 0 \cdot f(x')\|_Y = \|0 - 0\|_Y = 0 < \varepsilon$ proving the continuity of $\alpha \cdot f$

   b. $(\boldsymbol{\alpha \neq 0})$ by continuity there exists a $\delta \in \mathbb{R}_+$ such that if $\|x - x'\|_X < \delta$ then $|f(x) - f(x')| < \frac{\varepsilon}{|\alpha|}$ so that $|\alpha \cdot f(x) - \alpha \cdot f(x')| = |\alpha \cdot (f(x) - f(x'))| = |\alpha| \cdot |f(x) - f(x')| < |\alpha| \cdot \frac{\varepsilon}{|\alpha|} = \varepsilon$

Proving that $f + g$ and $\alpha \cdot f$ are continuous. □

Using inducion we can extend the above to a finite set of functions

Theorem 12.171. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces over $\mathbb{R}(\mathbb{C})$ and $\{f_i | X \rightarrow Y\}_{i \in \{1,\ldots,n\}}$ a finite family of continuous functions then $\sum_{i=1}^n f_i$: $X \rightarrow Y$ defined by $(\sum_{i=1}^n f_i)(x) = \sum_{i=1}^n f_i(x)$ is continuous.*

Proof. Let $\mathcal{S} = \{n \in \mathbb{N} |$let $\{f_i: X \rightarrow Y\}_{i \in \{1,\ldots,n\}}$ be a family continuous functions then $\sum_{i=1}^n f_i$ is continuous$\}$ then we have :

$\boldsymbol{1 \in \mathcal{S}}$. As $(\sum_{i=1}^1 f_i)(x) = \sum_{i=1}^1 f_i(x) = f_1(x)$ we have that $\sum_{i=1}^1 f_i = f_1$ which is continuous hence $1 \in \mathcal{S}$

$\boldsymbol{n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}}$. Let $\{f_i: X - Y\}_{i \in \{1,\ldots,n+1\}}$ a family of continuous functions then we have $(\sum_{i=1}^n f_i)(x) = \sum_{i=1}^{n+1} f_i(x) = f_{n+1}(x) + \sum_{i=1}^n f_i(x) = (f_{n+1} + \sum_{i=1}^n f_i)(x)$ proving that $\sum_{i=1}^{n+1} f_i = f_{n+1} + \sum_{i=1}^n f_i$ which as f_{n+1} is continuous and $\sum_{i=1}^n f_i$ is continuuos (because $n \in \mathcal{S}$) means by the previous theorem that $\sum_{i=1}^{n+1} f_i$ is continuous. Hence $n + 1 \in \mathcal{S}$. □

Theorem 12.172. *Let $\langle X, \|\|_X\rangle$ be a normed spaces over $\mathbb{R}(\mathbb{C})$ and $f: X \rightarrow \mathbb{R}(\mathbb{C})$, $g: X \rightarrow \mathbb{R}(\mathbb{C})$ be continuous functions $f \cdot g$ is continuous.*

Proof. Let $x \in X$and $\varepsilon \in \mathbb{R}_+$. By continuity of f, g there exists a $\delta_1 \in \mathbb{R}_+$ such that if $\|x - x'\|_X < \delta_1$ then $|f(x) - f(x')| < 1 \Rightarrow |f(x')| \leqslant |f(x') - f(x) + f(x)|_Y \leqslant |f(x') - f(x)| + |f(x)| \leqslant 1 + |f(x)|$. Also choose a δ_2 such that if $\|x - x'\| < \delta_3$ then $|f(x) - f(x')| < \frac{\varepsilon}{2(1 + |g(x)|)}$. Now choose $\delta_3 \in \mathbb{R}_+$ such that if $\|x - x'\|_X < \delta_3$ then $|g(x) - g(x')| < \frac{\varepsilon}{2(1 + |f(x)|)}$. Take then $\delta = \min(\delta_1, \delta_2, \delta_3)$ then if $\|x' - x\|_X < \delta$ we have

$$\begin{aligned} |f(x) \cdot g(x) - f(x') \cdot g(x')| &= |f(x) \cdot g(x) - f(x') \cdot g(x) + f(x') \cdot g(x) - f(x') \cdot g(x')| \\ &\leqslant |f(x) \cdot g(x) - f(x') \cdot g(x)\|_Y + \|f(x') \cdot g(x) - f(x') \cdot \\ &\quad g(x')| \\ &\leqslant |g(x)| \cdot |f(x) - f(x')| + |f(x')| \cdot |g(x) - g(x')| \\ &< |g(x)| \cdot \frac{\varepsilon}{2(1 + |g(x)|)} + (1 + |f(x)|) \cdot \frac{\varepsilon}{2(1 + |f(x)|)} \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

proving that $f \cdot g$ is continuous at x. As $x \in X$ was chosen arbitrary we have that $f \cdot g$ is continuous. □

Corollary 12.173. *Let $n\in\mathbb{N}_0$, $\langle X,\|\|_X\rangle$ be a normed spaces over $\mathbb{R}(\mathbb{C})$ and f: $X\to\mathbb{R}(\mathbb{C})$ then $f^n: X\to\mathbb{R}(\mathbb{C})$ defined by $f^n(x)=(f(x))^n$ is continuous*

Proof. Let $S=\{n\in\mathbb{N}_0|f^n: X\to\mathbb{R}(\mathbb{C})$ is a continuous function$\}$ then we have

- **$0\in\mathcal{S}$.** If $n=0$ then $f^0: X\to\mathbb{R}(\mathbb{C})$ is C_1 which is continuous by 12.136.
- **$0\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$.** As $\forall x\in X$ we have that $f^{n+1}(x)=(f(x))^{n+1}=f(x)\cdot(f(x))^n=(f\cdot f^n)(x)$ which is continuous because f^n is continuous as $n\in\mathcal{S}$, f is continuous and the above theorem. □

Using this corollary on the identity function 1_X proves the following corollary.

Corollary 12.174. *The mapping $.^n: \mathbb{R}\to\mathbb{R}$ defined by $x\to x^n$ where $n\in\mathbb{N}_0$ is continuous in the normed topology defined by $||$*

12.6 Linear maps and continuity

In this chapter we assume that all the normed spaces are defined over the same field $\mathbb{K}$ (which is either $\mathbb{R}$ or $\mathbb{C}$).

Theorem 12.175. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed vector spaces over the real (complex) field $\mathbb{K}$ and $L: X\to Y$ a linear mapping ($L\in\mathrm{Hom}(X,Y)$ then the following are equivalent*

1. *L is continuous*
2. *L is continuous at $0\in X$*
3. *$\exists k\in\mathbb{R}_+$ such that $\forall x\in X\vdash\|x\|_X=1$ we have $\|L(x)\|_Y\leqslant k$*
4. *$\exists k\in\mathbb{R}_+$ such that $\forall x\in X$ we have $\|L(x)\|_Y\leqslant k\cdot\|x\|_X$*

Proof.

1. **($1\Rightarrow 2$)** If L is continuous then by definition it is continuous at $0\in X$
2. **($2\Rightarrow 3$)** Take $1\in\mathbb{R}_+$ then using 12.151 there exists a $\delta\in\mathbb{R}_+$ such that if $x\in X$ and $\|x\|_X=\|0-x\|<\delta\Rightarrow\|L(x)\|_Y=\|L(x-0)\|_Y=\|L(x)-L(0)\|_Y=\|L(0)-L(x)\|_Y<1$. Now $\forall x\in X\vdash\|x\|_X=1$ take then $0<\delta'<\delta$ (see 9.57) then $\|\delta'\cdot x\|_X=|\delta'|\cdot\|x\|_X=\delta'\cdot 1=\delta'<\delta\Rightarrow\delta'\cdot\|L(x)\|_Y=|\delta'|\cdot\|L(x)\|_Y=\|\delta'\cdot L(x)\|_Y=\|L(\delta'\cdot x)\|_Y<1\underset{\delta'>0}{\Rightarrow}\|L(x)\|_Y<\frac{1}{\delta'}$. So if we take $k=\frac{1}{\delta'}$ then $\forall x\in X$ with $\|x\|_X=1$ we have $\|L(x)\|_Y<k\Rightarrow\|L(x)\|_Y<k$
3. **($3\Rightarrow 4$)** Let $k\in\mathbb{R}_+$ be such that $\forall x\in X$ with $\|x\|_X=1$ we have $\|L(x)\|_Y\leqslant k$. Take then $x\in X$ then we have the following possibilities
   a. **($\|x\|_X=0$)** then $x=0$ so that $L(x)=0\Rightarrow\|x\|_X=0\wedge\|L(x)\|_Y=0\Rightarrow\|x\|_X=0\leqslant 0=k\cdot 0=\|L(x)\|_X$
   b. **($\|x\|_X\neq 0$)** then we have that $\left\|\frac{1}{\|x\|_X}\cdot x\right\|_X=\left|\frac{1}{\|x\|_X}\right|\cdot\|x\|_X=\frac{1}{\|x\|_X}\cdot\|x\|_X=1$ so that $\frac{1}{\|x\|_X}\cdot\|L(x)\|_Y=\left|\frac{1}{\|x\|_X}\right|\cdot\|L(x)\|_Y=\left\|\frac{1}{\|x\|_X}\cdot L(x)\right\|_Y=\left\|L\left(\frac{1}{\|x\|_X}\cdot x\right)\right\|_Y\leqslant k\Rightarrow\|L(x)\|_Y\leqslant k\cdot\|x\|_X$

4. **(4 ⇒ 1)** Let $k\in\mathbb{R}_+$ such that $\forall x\in X$ we have $\|L(x)\|_Y\leqslant k\cdot\|x\|_X$. If now $x\in X$ and $\varepsilon>0$ take then $\delta=\frac{\varepsilon}{k}$ then if $y\in X$ with $\|x-y\|_X<\delta\Rightarrow\|L(x)-L(y)\|_Y=\|L(x-y)\|_Y\leqslant k\cdot\|x-y\|_Y<k\cdot\frac{\varepsilon}{k}=\varepsilon$ proving that L is continuous at x. As x is chosen arbitrary we have that L is continuous. □

Corollary 12.176. *Let $\langle\mathbb{R}^n,\|\|_n\rangle$, $\langle\mathbb{R}^m,\|\|_m\rangle$ be equipped with the maximum norm then every linear function $L:\mathbb{R}^n\to\mathbb{R}^m$ is continuous*

Proof. Let $\{\mathcal{E}_i\}_{i\in\{1,\ldots,n\}}$ be the canonical basis of $\mathbb{R}^n$ (see 10.161) then $\forall x\in\mathbb{R}^n$ we have that $x=(x_1,\ldots,x_n)=\sum_{i\in\{1,\ldots,n\}}x_i\cdot\mathcal{E}_i$ where $x_i=\pi_i(x)$ so that $\|L(x)\|_m=\left\|\sum_{i\in\{1,\ldots,n\}}\pi_i(x)\cdot\mathcal{E}_i\right\|_m\leqslant\sum_{i\in\{1,\ldots,n\}}\|\pi_i(x)\cdot\mathcal{E}_i\|=\sum_{i\in\{1,\ldots,n\}}|\pi_i(x)|\cdot\|\mathcal{E}_i\|\leqslant k'\cdot\sum_{i\in\{1,\ldots,n\}}|\pi_i(x)|$ where $k'=\max(\{\|L(\mathcal{E}_i)\|_m|i\in\{1,\ldots,n\}\})$. As $\max(\{|\pi_i(x)||i\in\{1,\ldots,n\}\})=\|x\|_n$ we have then $\|L(x)\|_m\leqslant k'\cdot n\cdot\|x\|_n$ or if we take $k=k'\cdot n\in\mathbb{R}_+$ we have $\|L(x)\|_m\leqslant k\cdot\|x\|_n$ proving continuity by the previous theorem. □

Corollary 12.177. *Let $\langle\mathbb{R}^n,\|\|_n\rangle$, $\langle\mathbb{R}^n,\|\|_n\rangle$ be equipped with the maximum norm then every linear isomorphism $L:\mathbb{R}^n\to\mathbb{R}^n$ is a homeomorphism (see 12.160)*

Proof. Using the previous corollary (see 12.176) we have that L is continuous. As $L^{-1}:\mathbb{R}^n\to\mathbb{R}^n$ is also a linear isomorphism we have that L^{-1} is also continuous. □

Theorem 12.178. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces and let $L:X\to Y$ be a continuous function then if we define $A_{L,s}=\{k\in\mathbb{R}_+|\forall x\in X\vdash\|x\|_X=1\text{ we have }\|L(x)\|_Y\leqslant k\}$, $A_{L,r}=\{k\in\mathbb{R}_+|\forall x\in X\vDash\|L(x)\|_Y\leqslant k\cdot\|x\|_X\}$ we have that*

1. $A_{L,s}=A_{L,r}$
2. $\inf(A_{L,s})=\inf(A_{L,r})$ *exists (is finite) and* $\inf(A_{L,s})=\inf(A_{L,r})\geqslant 0$

Proof.

1. If $k\in A_{L,s}$ then $\forall x\in X$ we have the following cases to consider

   a. **($\|x\|_X=0$)** then $x=0$ and thus $\|L(x)\|_Y=\|L(0)\|_Y=\|0\|_Y=0\leqslant k.0=k\cdot\|0\|_X=k\cdot\|x\|_X$

   b. **($\|x\|_X\neq 0$)** then $\left\|\frac{1}{\|x\|_X}\right\|_X=\left|\frac{1}{\|x\|_X}\right|\cdot\|x\|_X=\frac{1}{\|x\|_X}\cdot\|x\|_X=1$ so that as $k\in A_s$ we have that $\frac{1}{\|x\|_X}\cdot\|L(x)\|_Y=\left\|\frac{1}{\|x\|_X}\cdot L(x)\right\|_Y=\left\|L\left(\frac{1}{\|x\|_X}\right)\right\|_Y\leqslant k\Rightarrow\|L(x)\|_Y\leqslant k\cdot\|x\|_X$

   So in all cases we have $\|L(x)\|_Y\leqslant k\cdot\|x\|_X$ proving that $x\in A_{L,r}$.

2. If $k\in A_{L,r}$ then $\forall x\in X$ with $\|x\|_X=1$ we have that $\|L(x)\|_Y\leqslant k\cdot\|x\|_X=k\cdot 1=k$ so that $k\in A_{L,s}$

Using (1) and (2) we have proved that $A_{L,s}=A_{L,r}$.

As L is a continuous linear mapping we have by 12.175 the existence of a $k\in A_{L,r}$ so that $\emptyset\neq A_{L,r}=A_{L,s}$. As $\forall k\in A_{L,r}$ we have that $0<k$ we see that $A_{L,r},A_{L,s}$ is bounded below by 0. Using then the fact that the reals are conditionally complete (see 9.43) and 2.176 we have that $\inf(A_{L,r})=\inf(A_{L,s})$ exists and $\inf(A_{L,s})=\inf(A_{L,s})\geqslant 0$ (by the definition of a infinum). □

The above theorem means that the following definition makes sense

Definition 12.179. (norm of linear continuous function) *Let* $\langle X, \| \|_X \rangle$, $\langle Y, \| \|_Y \rangle$ *be normed spaces and let* $L: X \rightarrow Y$ *be a continuous function then* $\|L\| = \inf(A_{L,s}) = \inf(A_{L,r})$ *is called the operator norm.*

Theorem 12.180. *Let* $\langle X, \| \|_X \rangle$, $\langle Y, \| \|_Y \rangle$ *be normed spaces over the real (complex) field* $\mathbb{K}$ *and let* $L, L_1, L_2: X \rightarrow Y$ *be linear continuous functions then we have*

1. $\forall x \in X$ *we have* $\|L(x)\|_Y \leqslant \|L\| \cdot \|x\|_Y$
2. $L_1 + L_2$ *is linear and continuous and* $\|L_1 + L_2\| \leqslant \|L_1\| + \|L_2\|$
3. *If* $L = C_0$ *(*$C_0: X \rightarrow Y$ where $C_0(x) = 0$*) then* $\|L\| = 0$
4. $\forall \alpha \in \mathbb{K}$ *we have that* $\alpha \cdot L$ *is continuous and* $\|\alpha \cdot L\| = |\alpha| \cdot \|L\|$

Proof.

1. If $x \in X$ then we have to consider the following cases
   a. $\boldsymbol{(x = 0)}$ then $\|L(x)\|_Y = \|L(0)\|_Y = \|0\|_Y = 0 \leqslant \|L\| \cdot 0 = \|L\| \cdot \|0\|_Y = \|L\| \cdot \|x\|_X \Rightarrow \|L(x)\|_Y \leqslant \|L\| \cdot \|x\|_X$
   b. $\boldsymbol{(x \neq 0)}$ we proceed now by contradiction, so assume that $\|L\| \cdot \|x\|_X < \|L(x)\|_Y \Rightarrow \|L\| < \frac{\|L(x)\|_Y}{\|x\|_X}$ then by 2.167 there exists a $k \in A_{L,r} = \{k \in \mathbb{R}_+ | \forall x \in X \vDash \|L(x)\|_Y \leqslant k \cdot \|x\|_X\}$ such that $\|L\| \leqslant k < \frac{\|L(x)\|_Y}{\|x\|_X} \Rightarrow k \cdot \|x\| < \|L(x)\|_Y$ contradicting $k \in A_{L,r}$. So we must have that $\|L(x)\|_Y \leqslant \|L\| \cdot \|x\|_X$
2. Because the sum of continuous (linear) functions is continuous (linear) (see 12.170 and 10.185) we have that $L_1 + L_2$ is linear and continuous. Using (1) we have then that $\forall x \in X$ then $\|L_1(x)\|_Y \leqslant \|L_1\| \cdot \|x\|_X$ and $\|L_2(x)\|_Y \leqslant \|L_2\| \cdot \|x\|_X$ so that $\|(L_1 + L_2)(x)\|_Y = \|L_1(x) + L_2(x)\|_Y \leqslant \|L_1(x)\|_Y + \|L_2(x)\|_Y \leqslant \|L_1\| \cdot \|x\|_X + \|L_2\| \cdot \|x\|_X = (\|L_1\| + \|L_2\|) \cdot \|x\|_X$ so that $\|L_1\| + \|L_2\| \in A_{L_1+L_2,r} \Rightarrow \|L_1 + L_2\| = \inf(A_{L_1+L_2,r}) \leqslant \|L_1\| + \|L_2\|$
3. If $L = C_0$ then $\forall x \in X$ we have $L(x) = 0 \Rightarrow \|L(x)\|_Y = 0 \leqslant 0 = 0 \cdot_{\|x\|} X$ so that $0 \in A_{C_0,r}$ and thus $0 \leqslant \|C_0\| = \inf(A_{C_0,r}) \leqslant 0 \Rightarrow \|C_0\| = 0$
4. Using 12.170 and 10.185 we have that $\alpha \cdot L$ is a continuous linear mapping. Because of (1) we have $\forall x \in X$ that $\|L(x)\|_Y \leqslant \|L\| \cdot \|x\|_X$ so that $\|\alpha \cdot L(x)\|_Y = |\alpha| \cdot \|L(x)\|_Y \leqslant |\alpha| \cdot (\|L\| \cdot \|x\|_X) = (|\alpha| \cdot \|L\|) \cdot \|x\|_X$ so that $|\alpha| \cdot \|L\| \in A_{\alpha \cdot L,r} \Rightarrow \|\alpha \cdot L\| \leqslant |\alpha| \cdot \|L\|$. We consider now two cases for $\alpha \in \mathbb{K}$
   a. $\boldsymbol{(\alpha = 0)}$ then $\alpha \cdot L = 0 \cdot L = C_0$ and thus $\|\alpha \cdot L\| = \|C_0\| = 0 = |0| \cdot \|L\|$ so we have $\|\alpha \cdot L\| = |\alpha| \cdot \|L\|$
   b. $\boldsymbol{(\alpha \neq 0)}$ then $|\alpha| \neq 0$ then $\forall x \in X$ we have $\|\alpha \cdot L(x)\|_Y = |\alpha| \cdot \|L(x)\|_Y$ so that $\|L(x)\| = \frac{\|\alpha \cdot L(x)\|_Y}{|\alpha|} \leqslant \frac{\|\alpha \cdot L\| \cdot \|x\|_X}{|\alpha|}$ so that $\frac{\|\alpha \cdot L\|}{|\alpha|} \in A_{L,r} \Rightarrow \|L\| \leqslant \frac{\|\alpha \cdot L\|}{|\alpha|} \Rightarrow |\alpha| \cdot \|L\| \leqslant \|\alpha \cdot L\|$ this together with $\|\alpha \cdot L\| \leqslant |\alpha| \cdot \|L\|$ □

Theorem 12.181. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces over the real (complex) field $\mathbb{K}$, let $L: X \rightarrow Y$ be a continuous function and let X be non trivial (so there exists a $x \in X$ with $x \neq 0$) then we have if $A_{L,t} = \{\|L(x)\|_Y | x \in X \text{ with } \|x\|_X = 1\}$ and $A_{L,u} = \{\|L(x)\|_Y | x \in X \text{ with } \|x\|_X \leqslant 1\}$ that $\|L\| = \sup(A_{L,t}) = \sup(A_{L,u})$*

Proof. If $k \in A_{L,t}$ then $\exists x \in X$ with $\|x\|_X = 1 \leqslant 1$ such that $k = \|L(x)\|_Y \Rightarrow k \in A_{L,u}$ proving that

$$A_{L,t} \subseteq A_{L,u} \tag{12.18}$$

As $\exists x_0 \in X$ with $x_0 \neq 0 \Rightarrow \left\| \frac{1}{\|x\|_X} \cdot x_0 \right\|_X = 1 \leqslant 1$ so that $\frac{1}{\|x_0\|_X} \cdot x \in A_{L,t}, A_{L,u}$ and thus

$$\emptyset \neq A_{L,t} \wedge \emptyset \neq A_{L,u} \tag{12.19}$$

Now if $k \in A_{L,u}$ then $\forall x \in X$with $\|x\|_X \leqslant 1$ we have that $k = \|L(x)\|_Y$ and as $\|L(x)\|_Y \leqslant \|L\| \cdot \|x\|_X \leqslant \|L\|$ we see that $\|L\|$ is a upper bound of $A_{L,u}$ and by 12.18 also of $A_{L,t}$. Using 9.43 we have that $\sup(A_{L,u})$ and $\sup(A_{L,t})$ exists and that

$$\sup(A_{L,t}) \leqslant \sup(A_{L,u}) \leqslant \|L\| \tag{12.20}$$

(using also 12.18 and 2.171). Now $\forall x \in X$ with $\|x\|_X = 1$ we have $\|L(x)\|_Y \in A_{L,t} \Rightarrow \|L(x)\|_Y \leqslant \sup(A_{L,t}) \Rightarrow \sup(A_{L,t}) \in A_{L,s} \Rightarrow \|L\| = \inf(A_{L,s}) \leqslant \sup(A_{L,t})$ which together with 12.20 gives that $\|L\| = \sup(A_{L,t}) = \sup(A_{L,u})$ □

Example 12.182. If $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1,\ldots,n\}}, n \in \mathbb{N}$ be a family of normed spaces, $\left\langle \prod_{i \in \{1,\ldots,n\}} X_i, \|\| \right\rangle$ is the product space with the maximum norm then $\forall i \in \{1, \ldots, n\}$ we have that $\pi_i: \prod_{i \in \{1,\ldots,n\}} X_i \rightarrow X_i$ is a continuous linear mapping with $\|\pi_i\| < 1$.

Proof.

1. **(linearity)** I $x, y \in \prod_{i \in \{,\ldots,n\}}, \alpha, \beta \in \mathbb{K}$ then $\pi_i(\alpha \cdot x + \beta \cdot y) = (\alpha \cdot x + \beta \cdot y)_i = \alpha \cdot x_i + \beta \cdot y_i = \alpha \cdot \pi_i(x) + \beta \cdot \pi_i(y)$ proving that π_i is linear.
2. **(continuity)** If $x \in \prod_{i \in \{1,\ldots,n\}} X_i$ then $\|\pi_i(x)\|_i = \|x_i\|_i \leqslant \{\|x_i\|_i | i \in \{1, \ldots, n\}\} = \|x\| \leqslant 1 \cdot \|x\|$ proving that π_i is continuous with $\|\pi_i\| \leqslant 1$ as $\|\pi_i\| = \inf(\{k \in \mathbb{R}_+ | \forall x \in X \vDash \|\pi_i(x)\|_i \leqslant k \cdot \|x\|_i\}) \leqslant 1$ □

Theorem 12.183. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ and $\langle Z, \|\|_Z\rangle$ be normed spaces over the real (complex) field and $L_1: X \rightarrow Y$, $L_2: Y \rightarrow Z$ be linear continuous functions then $L_2 \circ L_1$ is linear and continuous and $\|L_2 \circ L_1\| \leqslant \|L_1\| \cdot \|L_2\|$*

Proof. By 12.143 and 10.182 we have that $L_2 \circ L_1$ is linear and continue. Now $\forall x \in X$ we have that $\|(L_2 \circ L_1)(x)\|_Z = \|L_2(L_1(x))\|_Z \leqslant \|L_2\| \cdot \|L_1(x)\|_Y \leqslant \|L_2\| \cdot \|L_1\| \cdot \|x\|_X$ meaning that $\|L_1\| \circ \|L_2\| \in A_{L_2 \circ L_1,,r}$ and thus that $\|L_2 \circ L_1\| = \inf(A_{L_2 \circ L_1}) \leqslant \|L_1\| \cdot \|L_2\|$ □

Definition 12.184. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces over the real (complex) field $\mathbb{K}$ then $L(X, Y) = \{L \in \mathrm{Hom}(X, Y) | L \,\mathrm{is\,continuous}\} \subseteq \mathrm{Hom}(X, Y)$ is the set of all linear continuous functions between X and Y.*

Theorem 12.185. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces over the real (complex) field $\mathbb{K}$ then $\langle L(X,Y),+,\cdot\rangle$ forms a subspace of* $\mathrm{Hom}(X,Y)$ *(and is thus a vector space). Furthermore if we define $\|\|: L(X,Y)\rightarrow\mathbb{R}_+$ by $L\rightarrow\|L\|$ then $\langle L(X,Y),\|\|\rangle$ forms a normed vector space over $\mathbb{K}$*

Proof. If $\alpha,\beta\in\mathbb{K}$ and $L_1,L_2\in L(X,Y)$ then by12.180 we have that $\alpha\cdot L_1+\beta\cdot L_2\in L(X,Y)$. Also using 12.180 we have that $\|\|$ is a pseudo norm. To prove that it is a norm assume that $\|L\|=0$ then $\forall x\in X$ we have $\|L(x)\|_Y\leqslant\|L\|\cdot\|x\|_X=0\cdot\|x\|_X=0\Rightarrow\|L(x)\|_Y=0\Rightarrow L(x)=0$ and thus $L=C_0=0$ (the zero function). □

As $\|L\|=\{\|L(x)\|_Y|x\in X \text{ with } \|x\|_X=1\}$ we see that $\|\|$ is dependent on the norm chosen for Y and X. We prove now that from the point of the generated topologies there is no difference.

Theorem 12.186. *Let X,Y be vector spaces over the field $\mathbb{K}$ and let $\|\|_{X_1},\|\|_{X_2}$ be two equivalent norms on X and let $\|\|_{Y_1}$, $\|\|_{Y_2}$ be two equivalent norms on Y. Then if $\|\|_1$ is a norm on $L(X,Y)$ based on $\|\|_{X_1},\|\|_{Y_1}$ and $\|\|_2$ is a norm based on $\|\|_{X_2}$, $\|\|_{Y_2}$ then $\|\|_1,\|\|_2$ are equivalent norms. So the topology of $L(X,Y)$ only depends on the topologies of X,Y*

Proof. First given $L\in L(X,Y)$ then if $A_{1,L}=\{k|\forall x\in X \text{ we have } \|L(x)\|_{Y_1}\leqslant k\cdot\|x\|_{X_1}\}$ $\|L\|_1=\inf(A_{1,L})$, if $A_{2,L}=\{k|\forall x\in X \text{ we have } \|L(x)\|_{Y_2}\leqslant k\cdot\|x\|_{X_2}\}$ then $\|L\|_2=\inf(A_{2,L})$. As $\|\|_{X_1},\|\|_{X_1}$ are equivalent and $\|\|_{Y_1},\|\|_{Y_2}$ are equivalent there exists (see 12.81) $\alpha_X>0,\alpha_Y>0,\beta_X>0,\beta_Y>0$ such that $\forall x\in X,y\in Y$ we have $\alpha_X\cdot\|x\|_{X_2}\leqslant\|x\|_{X_1}\leqslant\beta_X\cdot\|x\|_{X_2}$ and $\alpha_Y\cdot\|y\|_{Y_2}\leqslant\|y\|_{Y_1}\leqslant\beta_Y\cdot\|y\|_{Y_2}$. Then we have :

1. If $k\in A_{1,L}$ then $\forall x\in X$ we have $\|L(x)\|_{Y_1}\leqslant k\cdot\|x\|_{X_1}\Rightarrow\alpha_Y\cdot\|L(x)\|_{Y_2}\leqslant\|L(x)\|_{Y_1}\leqslant k\cdot\|x\|_{X_1}\leqslant(k\cdot\beta_X)\cdot\|x\|_{X_2}\Rightarrow\|L(x)\|_{Y_2}\leqslant\frac{k\cdot\beta_X}{\alpha_Y}\cdot\|x\|_{X_2}\Rightarrow\frac{k\cdot\beta_X}{\alpha_Y}\in A_{2,L}\Rightarrow\inf(A_{2,L})\leqslant k\cdot\frac{\beta_X}{\alpha_Y}\Rightarrow\frac{\alpha_Y}{\beta_X}\cdot\inf(A_{2,L})\leqslant k$. Assume now that $\inf(A_{1,L})<\frac{\alpha_Y}{\beta_X}\cdot\inf(A_{2,L})\Rightarrow\exists k\in A_{1,L}$ with $\inf(A_{1,L})\leqslant k<\frac{\alpha_Y}{\beta_X}\cdot\inf(A_{2,L})\leqslant k$ a contradiction so that $\frac{\alpha_Y}{\beta_X}\cdot\|L\|_{X_2}=\frac{\alpha_Y}{\beta_X}\cdot\inf(A_{2,L})\leqslant\inf(A_{1,L})=\|L\|_1$

2. If $k\in A_{2,L}$ then $\forall x\in X$ we have $\|L(x)\|_{Y_2}\leqslant k\cdot\|x\|_{X_2}\Rightarrow\frac{1}{\beta_Y}\cdot\|L(x)\|_{Y_1}\leqslant\|L(x)\|_{Y_2}\leqslant k\cdot\|x\|_{X_2}\leqslant\left(\frac{1}{\alpha_X}\cdot k\right)\cdot\|x\|_{X_1}\Rightarrow\|L(x)\|_{Y_1}\leqslant\frac{\beta_Y\cdot k}{\alpha_X}\cdot\|x\|_{X_1}\Rightarrow\frac{k\cdot\beta_Y}{\alpha_X}\in A_{1,L}\Rightarrow\inf(A_1,L)\leqslant k\cdot\frac{\beta_Y}{\alpha_X}\Rightarrow\frac{\alpha_X}{\beta_Y}\cdot\inf(A_{1,L})\leqslant k$. Assume now that $\inf(A_{2,L})<\frac{\alpha_X}{\beta_Y}\cdot\inf(A_{1,L})\Rightarrow\exists k\in A_{2,L}$ such that $\inf(A_{2,L})\leqslant k<\frac{\alpha_X}{\beta_Y}\cdot\inf(A_{1,L})\leqslant k$ a contradiction so that $\frac{\alpha_X}{\beta_Y}\cdot\inf(A_{1,L})\leqslant\inf(A_{2,L})\Rightarrow\inf(A_{1,L})\leqslant\frac{\beta_Y}{\alpha_X}\cdot\inf(A_{2,L})\Rightarrow\|L\|_1\leqslant\frac{\beta_Y}{\alpha_X}\cdot\|L\|_2$

Using (1) and (2) we have $\frac{\alpha_Y}{\beta_X}\cdot\|L\|_2\leqslant\|L_1\|\leqslant\frac{\beta_Y}{\alpha_X}\cdot\|L\|_2$ which as $0<\frac{\alpha_Y}{\beta_X},\frac{\beta_Y}{\alpha_X}$ and L is chosen arbitrary proves that $\|\|_1$ and $\|\|_2$ are equivalent. □

Theorem 12.187. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed vector spaces on $\mathbb{K}$ then a linear function $L: X\rightarrow Y$ is open if and only if there exists $\delta>0$ such that $B_{\|\|_Y}(0,\delta)\subseteq L(B_{\|\|_X}(0,1))$*

Proof.

1. **(⇒)** This trivial using 12.149

2. **(⇐)** Using 12.149 we only have to prove that $\forall x \in X$ and $\delta > 0$ there exists a $\delta' > 0$ such that $L(x) \subseteq B_{\|\|_Y}(L(x), \delta') \subseteq L(B_{\|\|_X}(x, \delta))$. Now as we now by the hypothesis that there exists a $\delta'' > 0$ such that $B_{\|\|_Y}(0, \delta'') \subseteq L(B_{\|\|_X}(0, 1))$ we have:

$$
\begin{aligned}
y \in B_{\|\|_Y}\left(L(x), \frac{\delta''}{\delta}\right) &\Rightarrow \|y - L(x)\|_Y < \frac{\delta''}{\delta}\\
&\Rightarrow \|\delta \cdot (y - L(x))\|_Y < \delta''\\
&\Rightarrow \delta \cdot (y - L(x)) \in B_{\|\|_Y}(0, \delta'')\\
&\Rightarrow \delta \cdot (y - L(x)) \in L(B_{\|\|_X}(0, 1))\\
&\Rightarrow \exists x_0 \in X \text{ with } \|x_0\|_X < 1 \text{ such that } L(x_0) = \delta \cdot (y - L(x))\\
&\Rightarrow y = \frac{L(x_0)}{\delta} + L(x) \text{ take } x_1 = \frac{x_0}{\delta}\\
&\Rightarrow \|x_1\| < \delta \wedge y = L(x_1) + L(x) = L(x_1 + x) \text{ take now } x_2 = x_1 + x\\
&\Rightarrow \|x_2 - x\| < \delta \text{ and } y = L(x_2)\\
&\Rightarrow y \in L(B_{\|\|_Y}(x, \delta))
\end{aligned}
$$

□

12.7 Multilinear mappings and continuity

In this chapter we assume that all the normed spaces are defined over the same field $\mathbb{K}$ (which is either $\mathbb{R}$ or $\mathbb{C}$).

Theorem 12.188. *Let $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1,\ldots,n\}}$ be normed vector spaces over the real (complex) field $\mathbb{K}$, $\langle Y, \|\|_Y\rangle$ a normed vector space over the field $\mathbb{K}$ and L: $\prod_{i \in \{1,\ldots,n\}} X_i \rightarrow Y$ a multi-linear function then we have the following equivalences (using the product topology on $\prod_{i \in \{1,\ldots,n\}} X_i$) and the norm topology on Y.*

1. *L is continuous*
2. *L is continuous at $0 \in \prod_{i \in \{1,\ldots,n\}} X_i$*
3. *$\exists k \in \mathbb{R}_+$ such that $\forall x \in \prod_{i \in \{1,\ldots,n\}} X_i$ with $\forall i \in \{1,\ldots,n\}$ we have that $\|x_i\|_i = 1$ then $\|L(x)\|_Y \leqslant k$*
4. *$\exists k \in \mathbb{R}_+$ such that $\forall x \in \prod_{i \in \{1,\ldots,n\}} X_i$ we have $\|L(x)\|_Y \leqslant k \cdot \prod_{i \in \{1,\ldots,n\}} \|x_i\|_i$*

Proof. First note that by 12.79 then product topology is generated by the norm $\|x\| = \max(\|\pi_i(x)\|_i | i \in \{1,\ldots,n\}) = \max(\|x_i\|_i | i \in \{1,\ldots,n\})$

1. **(1 ⇒ 2)** this is trivial

2. **(2 ⇒ 3)** using 12.151 given $\varepsilon = 1$ the exists a $\delta \in \mathbb{R}_+$ such that $\forall x \in \prod_{i \in \{1, \ldots, \mathrm{nn}\}} X_i$ with $\|x\| = \|x - 0\| < 1$ we have $\|L(x)\|_Y = \|L(x) - L(0)\|_Y < 1$. As $0 < \delta$ there exists a $0 < \delta' < \delta$ take then $k = \frac{1}{\delta'}$. Then $\forall x \in \prod_{i \in \{1, \ldots, n\}} X_i$ with $\|x\| = \max\{\|x_i\| | i \in \{1, \ldots, n\}\} = 1$ then $\|\delta' \cdot x\| = |\delta'| \cdot \|x\| = \delta' \cdot \|x\| < \delta \cdot 1 = \delta \Rightarrow \delta' \cdot \|L(x)\|_Y = |\delta'| \cdot \|L(x)\|_Y = \|\delta' \cdot L(x)\|_Y = \|L(\delta' \cdot x)\|_Y < 1 \Rightarrow \|L(x)\|_Y < \frac{1}{\delta'} = k$. So $\forall x \in \prod_{i \in \{1, \ldots, n\}} X_i$ with $\|x\| = 1$ we have $\|L(x)\|_Y \leqslant k$

3. **(3 ⇒ 4)** Let $k \in \mathbb{R}_+$ such that $\forall x \in \prod_{i \in \{1, \ldots, n\}} X_i$ with $\|x\| = 1$ we have $\|L(x)\|_Y \leqslant k$ then $\forall x \in \prod_{i \in \{1, \ldots, n\}} X_I$ we have to consider the following cases

   a. **($\exists i \in \{1, \ldots, n\} \vdash x_i = 0$)** then by 10.240 we have $L(x) = 0 \Rightarrow \|L(x)\|_Y = 0 \leqslant k \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i$

   b. **($\forall i \in \{1, \ldots, n\} \vDash x_i \neq 1$)** define then $y = (y_1, \ldots, y_n)$ by $y_i = \frac{1}{\|x_i\|_i}$ then $\|y_i\|_i = 1$ and thus $\frac{1}{\prod_{i \in \{1, \ldots, n\}} \|x_i\|_i} \cdot \|L(x_1, \ldots, x_n)\|_Y \underset{10.56}{=} \left(\prod_{i \in \{1, \ldots, n\}} \frac{1}{\|x_i\|_i}\right) \cdot \|L(x_1, \ldots, x_n)\|_Y = \left\|\prod_{i \in \{1, \ldots, n\}} \frac{1}{\|x_i\|_i} \cdot L(x_1, \ldots, x_n)\right\|_Y = \left\|L\left(\frac{1}{\|x_1\|_1} \cdot x_1, \ldots, \frac{1}{\|x_n\|_n} \cdot x_n\right)\right\|_Y = \|L(y)\|_Y \leqslant k \Rightarrow \|L(x)\|_Y \leqslant k \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i$

4. **(4 ⇒ 1)** Let $x \in \prod_{i \in \{1, \ldots, n\}} X_i$ and $\varepsilon \in \mathbb{R}_+$ take then $0 < \delta < \min\left(1, \frac{\varepsilon}{k}\right)$ (k from the condition in (4)) then if $\|x - y\| < \delta$ then we have $\forall i \in \{1, \ldots, n\}$ we have $\|x_i - y_i\|_i = \|(x - y)_i\|_i < \delta$ so that $\|L(x) - L(y)\|_Y = \|L(x - y)\|_Y \leqslant k \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i - y_i\|_i < k \cdot \delta \leqslant \varepsilon$ [as $\forall i \in \{1, \ldots, n\}$ we have $\|x_i - y_i\| < \delta < 1$ and 10.64] proving continuity. □

Theorem 12.189. *Let $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1, \ldots, n\}}$ be normed vector spaces over the real (complex) field $\mathbb{K}$, $\langle Y, \|\|_Y\rangle$ a normed vector space over the field $\mathbb{K}$ and L: $\prod_{i \in \{1, \ldots, n\}} X_i \rightarrow Y$ a multi-linear function then if we define*

$$M_{L,s} = \left\{k \in \mathbb{R}_+ | \forall x \in \prod_{i \in \{1, \ldots, n\}} X_i \vdash \forall i \in \{1, \ldots, n\} \text{ we have } \|x_i\| = 1 \text{ then we have } \|L(x)\|_Y \leqslant k\right\}$$

$$M_{L,r} = \left\{k \in \mathbb{R}_+ | \forall x \in \prod_{i \in \{1, \ldots, n\}} X_i \text{ we have } \|L(x)\|_Y \leqslant k \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i\right\}$$

then $M_{L,s} = M_{L,r}$ and $\inf(M_{L,s}) = \inf(M_{L,r})$ exists and is finite (and obviously $\geqslant 0$)

Proof. First prove that $M_{L,s} = M_{L,r}$

1. If $k \in M_{L,s}$ and $x \in \prod_{i \in \{1, \ldots, n\}} X_i$ then we have two cases :

   a. **($\exists i \in \{1, \ldots, n\} \vdash \|x_i\|_i = 0$)** then $x_i = 0$ and thus using 10.240 we have that $L(x) = 0 \Rightarrow \|L(x)\|_Y = 0 = k \cdot 0 \underset{10.56}{=} k \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i \Rightarrow k \in M_{L,r}$

b. $(\forall \boldsymbol{i} \in \{\mathbf{1}, \ldots, \boldsymbol{n}\} \vDash \|\boldsymbol{x_i}\|_{\boldsymbol{i}} \neq \mathbf{0})$ then if we define $y \in \prod_{i \in \{1, \ldots, n\}} X_i$ by $y_i = \frac{1}{\|x_i\|_i}$ then $\|y_i\|_i = 1$ and thus $\frac{1}{\prod_{i \in \{1, \ldots, n\}} \|x_i\|_i} \cdot \|L(x)\|_Y \underset{10.56}{=} \left(\prod_{i \in \{1, \ldots, n\}} \frac{1}{\|x_i\|_i}\right) \cdot \|L(x)\|_Y = \left\|\left(\prod_{i \in \{1, \ldots, n\}} \frac{1}{\|x_i\|_i}\right) \cdot L(x)\right\|_Y = \|L(y)\|_Y \leqslant k \Rightarrow \|L(x)\|_Y \leqslant k \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i \Rightarrow k \in M_{L,r}$

2. If $k \in M_{L,r}$ and $x \in \prod_{i \in \{1, \ldots, n\}} X_i$ with $\|x_i\|_i = 1 \forall i \in I$ then $\|L(x)\|_Y \leqslant k \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i = k \cdot 1 = k \Rightarrow k \in M_{L,s}$

Next by 12.188 there exists a $k \in \mathbb{R}_+$ such that $\forall x \in \Pi_{i \in \{1, \ldots, n\}} X_i$ we have $\|L(x)\|_Y \leqslant k \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i$ and thus $k \in M_{L,s} \Rightarrow M_{L,s} \neq \emptyset$. Using then the fact that the reals are conditionally complete (see 9.43) and 2.176 we have that $\inf(M_{L,r}) = \inf(M_{L,s})$ exists and $\inf(A_{L,s}) = \inf(A_{L,s}) \geqslant 0$ (by the definition of a infinum). □

Definition 12.190. *Let $\{\langle X_i, \|\|_i \rangle\}_{i \in \{1, \ldots, n\}}$ be normed vector spaces over the real (complex) field $\mathbb{K}$, $\langle Y, \|\|_Y \rangle$ a normed vector space over the field $\mathbb{K}$ and $L: \prod_{i \in \{1, \ldots, n\}} X_i \rightarrow Y$ a multi-linear function then we define $\|L\| = \inf(M_{L,r}) = \inf(M_{L,s})$ to be the operator norm.*

Theorem 12.191. *Let $\{\langle X_i, \|\|_i \rangle\}_{i \in \{1, \ldots, n\}}$ be normed vector spaces over the real (complex) field $\mathbb{K}$, $\langle Y, \|\|_Y \rangle$ a normed vector space over the field $\mathbb{K}$ and L, L_1, L_2: $\prod_{i \in \{1, \ldots, n\}} X_i \rightarrow Y$ a multi-linear function then we have*

1. *$\forall x \in \prod_{i \in \{1, \ldots, n\}} X_i$ we have $\|L(x)\|_Y \leqslant \|L\| \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i$*
2. *If $\forall x \in \prod_{i \in \{1, \ldots, n\}} X_i$ we have $\|L(x)\|_Y \leqslant A \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i$ then $\|L\| \leqslant A$*
3. *If $\forall x \in \prod_{i \in \{1, \ldots, n\}} X_i$ with $\|x_i\| = 1$ we have $\|L(x)\|_Y \leqslant A$ then $\|L\| \leqslant A$*
4. *$L_1 + L_2$ is a continuous multilinear mapping and $\|L_1 + L_2\| \leqslant \|L_1\| + \|L_2\|$*
5. *If $L = C_0$ (the constant 0 function) then $\|L\| = 0$*
6. *If $\|L\| = 0$ then $L = C_0$*
7. *$\forall \alpha \in \mathbb{K}$ we have that $\alpha \cdot L$ is a continuous multilinear mapping and $\|\alpha \cdot L\| = |\alpha| \cdot \|L\|$*

Proof.

1. Let $x \in \prod_{i \in \{1, \ldots, n\}} X_i$ then we have the following possibilities

   a. $(\exists \boldsymbol{i} \in \{\mathbf{1}, \ldots, \boldsymbol{n}\} \vdash \boldsymbol{x_i} = \mathbf{0})$ then by 10.240 we have $L(x) = 0 \Rightarrow \|L(x)\|_Y = 0 \leqslant \|L\| \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i$

   b. $(\forall \boldsymbol{i} \in \{\mathbf{1}, \ldots, \boldsymbol{n}\} \vDash \boldsymbol{x_i} \neq \mathbf{0})$ then $\prod_{i \in \{1, \ldots, n\}} \|x_i\|_i > 0$, assume now that $\exists x$ so that $\|L\| \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\| < \|L(x)\|_Y \Rightarrow \|L\| < \frac{1}{\prod_{i \in \{1, \ldots, n\}} \|x_i\|_i} \cdot \|L(x)\|_Y$ as $\|L\| = \inf(M_{L,r})$ there exists a $k \in M_{L,r}$ such that $\|L\| \leqslant k < \frac{1}{\prod_{i \in \{1, \ldots, n\}} \|x_i\|_i} \cdot \|L(x)\|_Y \leqslant \frac{1}{\prod_{i \in \{1, \ldots, n\}} \|x_i\|_i} \cdot k \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i = k$ giving the contradiction $k < k$. We conclude thus that $\forall x \in \prod_{i \in \{1, \ldots, n\}} x_i$ with $x_i \neq 0 \forall i \in \{1, \ldots, n\}$ we have $\|L(x)\|_Y \leqslant \|L\| \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i$

2. This follows from the fact that $\|L\| = \inf(M_{L,r}) = \inf\left(\left\{k \in \mathbb{R}_+ | \forall x \in \prod_{i \in \{1,\ldots,n\}} X_i \text{ we have } \|L(x)\|_Y \leqslant k \cdot \prod_{i \in \{1,\ldots,n\}} \|x_i\|_i\right\}\right)$
3. This followed from the fact $\|L\| = \inf(M_{L,s}) = \inf\left(k \in \mathbb{R}_+ | \forall x \in \prod_{i \in \{1,\ldots,n\}} X_i \vdash \forall i \in \{1,\ldots,n\} \text{ we have } \|x_i\| = 1 \text{ then we have } \|L(x)\|_Y \leqslant k\right)$
4. Using 10.233 we have that $L_1 + L_2$ is multilinear and using 12.170 we have that $L_1 + L_2$ is continuous. If now $x \in \prod_{i \in \{1,\ldots,n\}} X_i$ then $\|L_1(x) + L_2(x)\|_Y \leqslant \|L_1(x)\|_Y + \|L_2(x)\|_Y \leqslant \|L_1\| \cdot \prod_{i \in \{1,\ldots,n\}} \|x_i\|_i + \|L_2\| \cdot \prod_{i \in \{1,\ldots,n\}} \|x_i\|_i = (\|L_1\| + \|L_2\|) \cdot \prod_{i \in \{1,\ldots,n\}} \|x_i\|_i \Rightarrow \|L_1\| + \|L_2\| \in M_{L_1+L_2,r} \Rightarrow \|L_1 + L_2\| = \inf(M_{L_1+L_2,r}) \leqslant \|L_1\| + \|L_2\|$
5. If $L = C_0$ then $\forall x \in \prod_{i \in \{1,\ldots,n\}} X_i$ that $\|L(x)\|_Y = \|C_0(x)\|_Y = \|0\|_Y = 0 \leqslant 0 \cdot \prod_{i \in \{1,\ldots,n\}} \|x_i\|_i$ proving that $0 \in M_{L,r} \Rightarrow 0 \leqslant \|L\| = \inf(M_{L,r}) \leqslant 0 \Rightarrow \|L\| = 0$
6. If $\|L\| = 0$ then if $x \in \prod_{i \in \{1,\ldots,n\}} X_i$ we have $\|L(x)\|_Y = 0 \Rightarrow L(x) = 0 \Rightarrow L = C_0$
7. Using 10.233 we have that $\alpha \cdot L$ is multilinear and using 12.170 we have that $\alpha \cdot L$ is continuous. If $x \in \prod_{i \in \{1,\ldots,n\}} X_i$ then $\|\alpha \cdot L(x)\|_Y = |\alpha| \cdot \|L(x)\|_Y \leqslant |\alpha| \cdot \|L\| \cdot \prod_{i \in \{1,\ldots,n\}} \|x_i\|_i \Rightarrow |\alpha| \cdot \|L\| \in M_{\alpha \cdot L,r} \Rightarrow \|\alpha \cdot L\| = \inf(M_{\alpha \cdot L,r}) \leqslant |\alpha| \cdot \|L\|$ so
$$\|\alpha \cdot L\| \leqslant |\alpha| \cdot \|L\| \tag{12.21}$$
Assume now that $\|\alpha \cdot L\| < |\alpha| \cdot \|L\|$ then consider the cases
   a. $(\boldsymbol{\alpha = 0})$ then $\alpha \cdot L = C_0$ and thus $0 = 0 \cdot \|L\| = |\alpha| \cdot \|L\| > \|\alpha \cdot L\| = \|C_0\| \underset{(3)}{=} 0 \Rightarrow 0 > 0$ a contradiction.
   b. $(\boldsymbol{\alpha \neq 0})$ then $|\alpha| \neq 0$ so that $\|\alpha \cdot L\| < |\alpha| \|L\|$ as $\|\alpha L\| = \inf(A_{\alpha \cdot L,r})$ there exist a $k \in A_{\alpha \cdot L,r}$ such that $\|\alpha \cdot L\| \leqslant k < |\alpha| \cdot \|L\|$ so if $x \in \prod_{i \in \{1,\ldots,n\}} X_i$ then $|\alpha| \cdot \|L(x)\|_Y = \|\alpha \cdot L(x)\|_Y \leqslant k \cdot \prod_{i \in \{1,\ldots,n\}} \|x_i\|_i \Rightarrow \|L(x)\|_Y \leqslant \frac{k}{|\alpha|} \cdot \prod_{i \in \{1,\ldots,n\}} \|x_i\|_i \Rightarrow \frac{k}{|\alpha|} \in M_{L,r} \Rightarrow \|L\| \leqslant \frac{k}{|\alpha|} < \frac{|\alpha| \cdot \|L\|}{|\alpha|} = \|L\|$ giving the contradiction $\|L\| < \|L\|$.

   as the assumption turns up a contradiction in all cases we must have $|\alpha| \cdot \|L\| \leqslant \|\alpha \cdot L\|$ which together with 12.21 means that $\|\alpha \cdot L\| = |\alpha| \cdot \|L\|$ □

From the above theorem we have then trivially the following theorem

Theorem 12.192. *Let $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1,\ldots,n\}}$ be normed vector spaces over the real (complex) field $\mathbb{K}$, $\langle Y, \|\|_Y\rangle$ a normed vector space over the real (complex) field $\mathbb{K}$ then if $L(X_1,\ldots,X_n;Y) = \{L \in \mathrm{Hom}(X_1,\ldots,X_n;Y) | L \text{ is continuous}\} \subseteq \mathrm{Hom}(X_1,\ldots,X_n;Y)$ [The set of multilinear mappings between $\prod_{i \in \{1,\ldots,n\}} X_i$ and Y] then $L(X_1,\ldots,X_n;Y)$ is a subspace of $\mathrm{Hom}(X_1,\ldots,X_n;Y)$ and $\langle L(X_1,\ldots,X_n;Y), \|\|\rangle$ ($\|\|$ the operator norm) forms a normed vector space.*

Theorem 12.193. *Let $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1,\ldots,n\}}$ be normed vector spaces over the real (complex) field $\mathbb{K}$, $\langle Y, \|\|_Y\rangle$, $\langle Z, \|\|_Z\rangle$ normed vector spaces over the real (complex) field $\mathbb{K}$ then we have if $L \in L(Y,Z)$ and $K \in L(X_1,\ldots,X_n;Y)$ then $L \circ K \in L(X_1,\ldots,X_n;Z)$ with $\|L \circ K\|_{L(X_1,\ldots,X_n;Z)} \leqslant \|L\|_{L(Y,Z)} \cdot \|K\|_{L(X_1,\ldots,X_n;Y)}$*

Proof. First multilinearity follows from 10.234, continuity follows from the fact that the composition of continuous functions is continuous. Finally if $x=(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}X_i$ then $\|(L\circ K)(x_1,\ldots,x_n)\|_Z=\|L(K(x_1,\ldots,x_n))\|_Z \leqslant_{L\,\text{is linear}} \|L\|_{L(X_1,\ldots,X_n;Y)}\cdot\|K(x_1,\ldots,x_n)\|_Y \leqslant_{K\,\text{is multilinear}} \|L\|_{L(Y,Z)}\cdot\|K\|_{L(X_1,\ldots,X_n;Y)}\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i \underset{12.191}{\Rightarrow} \|L\circ K\|_{L(X_1,\ldots,X_n;Z)}\leqslant\|L\|_{L(Y,Z)}\cdot\|K\|_{L(X_1,\ldots,X_n;Y)}$ □

We prove now that the operator norm depends only on the topology of $\langle X_i,\|\|_i\rangle$ and $\langle Y,\|\|_Y\rangle$ in the following theorem.

Theorem 12.194. *Let $\{\langle X_i,\|\|_{1,i}\rangle\}_{i\in\{1,\ldots,n\}}$, $\{\langle X_i,\|\|_{2,i}\rangle\}_{i\in\{1,\ldots,n\}}$, $\langle Y,\|\|_{1,Y}\rangle$, $\langle Y,\|\|_{2,Y}\rangle$ be normed vector spaces over the real (complex) field $\mathbb{K}$, with $\forall i\in\{1,\ldots,n\}$ $\|\|_{1,i}$ is equivalent with $\|\|_{2,i}$ and $\|\|_{1,Y}$ are equivalent then if $\|\|_1$ is the operator norm based on $\{\langle X_i,\|\|_{1,i}\rangle\}$, $\langle Y,\|\|_1\rangle$ and $\|\|_2$ is the operator norm based on $\{\langle X_i,\|\|_{2,i}\rangle\}$, $\langle Y,\|\|_2\rangle$ then $\|\|_1$ and $\|\|_2$ are equivalent.*

Proof. Given $L\in L(X_1,\ldots,X_n;Y)$ let $A_{1,L}=\left\{k\in\mathbb{R}_+|\forall x\in\prod_{i\in\{1,\ldots,n\}}X_i\text{ we have }\|L(x)\|_{Y_1}\leqslant k\cdot\prod_{i\in\{1,\ldots,\}}\|x_i\|_{1,i}\right\}$ and $A_{2,L}=\left\{k\in\mathbb{R}_+|\forall x\in\prod_{i\in\{1,\ldots,n\}}X_i\text{ we have }\|L(x)\|_{Y_2}\leqslant k\cdot\prod_{i\in\{1,\ldots,\}}\|x_i\|_{2,i}\right\}$ then $\|L\|_1=\inf(A_1,L)$ and $\|L\|_2=\inf(A_{2,L})$. By the hypothesis and the definition of equivalence of norms (see 12.82) there exists $\{(\alpha_i,\beta_i)\}_{i\in\{1,\ldots,n\}}$ and $\alpha,\beta>0$ such that $\forall i\in\{1,\ldots,n\}$ we have $\alpha_i,\beta_i>0$ and $\forall x\in X_i$ that $\alpha_i\cdot\|x\|_{2,i}\leqslant\|x\|_{1,i}\leqslant\beta_i\cdot\|x\|_{1,i}$. Also $\forall y\in Y$ we have $\alpha\cdot\|y\|_{2,Y}\leqslant\|y\|_{1,Y}\leqslant\beta\cdot\|y\|_{2,Y}$. Then we have

1. $\forall k\in A_{1,L}$ we have $\forall x\in\prod_{i\in\{1,\ldots,n\}}X_i$ that $\|L(x)\|_{1,Y}\leqslant k\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_{1,i} \underset{\beta_i>0,\|\|_{1,i}\geqslant 0}{\Rightarrow} \alpha\cdot\|L(x)\|_{2,Y}\leqslant\|L(x)\|_1\leqslant k\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_{1,i}\leqslant\left(k\cdot\prod_{i\in\{1,\ldots,n\}}\beta_i\right)\cdot\prod_{i\in\{1,\ldots,n\}}\|x\|_{2,i}\underset{\alpha>0}{\Rightarrow}\|L(x)\|_{2,Y}\leqslant\left(\frac{\prod_{i\in\{1,\ldots,n\}}\beta_i}{\alpha}\cdot k\right)\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_{2,i}$ proving that $\left(\frac{\prod_{i\in\{1,\ldots,n\}}\beta_i}{\alpha}\cdot k\right)\in A_{2,L}$ so that $\inf(A_{2,L})\leqslant\left(\frac{\prod_{i\in\{1,\ldots,n\}}\beta_i}{\alpha}\cdot k\right)\Rightarrow\frac{\alpha}{\prod_{i\in\{1,\ldots,n\}}\beta_i}\cdot\inf(A_{2,L})\leqslant k$. Assume that $\inf(A_{1,L})<\frac{\alpha}{\prod_{i\in\{1,\ldots,n\}}\beta_i}\cdot\inf(A_{2,L})$ then $\exists k\in A_{1,L}$ such that $\inf(A_{1,L})\leqslant k<\frac{\alpha}{\prod_{i\in\{1,\ldots,n\}}\beta_i}\cdot\inf(A_{2,L})\leqslant k$ a contradictions so that $\frac{\alpha}{\prod_{i\in\{1,\ldots,n\}}\beta_i}\cdot\|L\|_2=\frac{\alpha}{\prod_{i\in\{1,\ldots,n\}}\beta_i}\cdot\inf(A_{2,L})\leqslant\inf(A_{1,L})=\|L\|_1$

2. $\forall k\in A_{2,L}$ we have $\forall x\in\prod_{i\in\{1,\ldots,n\}}X_i$ that $\|L(x)\|_{2,Y}\leqslant k\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_{2,i}\Rightarrow\frac{1}{\beta}\cdot\|L(x)\|_{1,Y}\leqslant\|L(x)\|_{2,Y}\leqslant k\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_{2,y}\leqslant\frac{k}{\prod_{i\in\{1,\ldots,n\}}\alpha_i}\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_{1,i}\Rightarrow\|L(x)\|_{1,Y}\leqslant\left(\frac{\beta}{\prod_{i\in\{1,\ldots,n\}}\alpha_i}\right)\cdot k\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_{1,i}$ proving that $k\cdot\frac{\beta}{\prod_{i\in\{1,\ldots,n\}}\alpha_i}\in A_{1,L}$ so that $\inf(A_{1,L})\leqslant k\cdot\frac{\beta}{\prod_{i\in\{1,\ldots,n\}}\alpha_i}\Rightarrow\frac{\prod_{i\in\{1,\ldots,n\}}\alpha_i}{\beta}\cdot\inf(A_{1,L})\leqslant k$. Assume that $\inf(A_{2,L})<\frac{\prod_{i\in\{1,\ldots,n\}}\alpha_i}{\beta}\cdot\inf(A_{1,L})\Rightarrow\exists k\in A_{2,L}$ such that $\inf(A_{2,L})\leqslant k<\frac{\prod_{i\in\{1,\ldots,n\}}\alpha_i}{\beta}\cdot\inf(A_{1,L})\leqslant k$ a contradiction so that $\frac{\prod_{i\in\{1,\ldots,n\}}\alpha_i}{\beta}\cdot\|L\|_1=\frac{\prod_{i\in\{1,\ldots,n\}}\alpha_i}{\beta}\cdot\inf(A_{1,L})\leqslant\inf(A_{2,L})=\|L\|_2$

From (1),(2) and the fact that $\frac{\alpha}{\prod_{i\in\{1,\ldots,n\}}\beta_i}>0$, $\frac{\beta}{\prod_{i\in\{1,\ldots,n\}}\alpha_i}>0$ we have $\frac{\alpha}{\prod_{i\in\{1,\ldots,n\}}\beta_i}\cdot\|L\|_2\leqslant\|L\|_1\leqslant\frac{\beta}{\prod_{i\in\{1,\ldots,n\}}\alpha_i}\cdot\|L\|_2$ proving the equivalency of $\|\|_1$ and $\|\|_2$. □

Notation 12.195. *If $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}, n\in\mathbb{N}$ be family of normed spaces, where $\forall i\in\{1,\ldots,n\}$ we have $\langle X_i,\|\|_i\rangle=\langle X,\|\|_X\rangle$ a normed space, $\langle Y,\|\|_Y\rangle$ then $\langle L(X_1,\ldots,X_n;Y),\|\|\rangle$ is noted as $\langle L(X^n;Y),\|\|\rangle$ where $L\in L(X^n;Y)\Rightarrow L{:}\, X^n=\prod_{i\in\{1,\ldots,n\}}X\rightarrow Y$ is multilinear and continuous and has norm $\|L\|$. Note that $L(X^1;Y)=L(X,Y)$ by definition.*

Theorem 12.196. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ and $\langle Z,\|Z\|_Z\rangle$ be normed spaces over the real (complex) field $\mathbb{K}$ then the composition function $\circ{:}\, L(Y,Z)\times L(X,Y)\rightarrow L(X,Z)$ defined by $\circ(L_2,L_1)=L_2\circ L_1$ is multilinear and continuous (using the canonical norms) with $\|\circ\|=1$. So $\circ\in L(L(Y,Z),L(X,Y);L(X,Z))$*

Proof. First note that the composition of continuous linear mappings is again linear and continuous so that if $(L_2,L_1)\in L(Y,Z)\times L(X,Y)$ then $\circ(L_2,L_1)=L_2\circ L_1\in L(X,Z)$ proving that $\circ$ is indeed a mapping from $L(Y,Z)\times L(X,Y)$ to $L(X,Z)$.

To prove multilinearity note:

1. Let $\alpha,\beta\in\mathbb{K}$, $L_1,L_2\in L(Y,Z)$ and $L\in L(X,Y)$ then $\forall x\in X$ we have $\circ(\alpha\cdot L_1+\beta\cdot L_2,L)(x)=((\alpha\cdot L_1+\beta\cdot L_2)\circ L)(x)=(\alpha\cdot L_1+\beta\cdot L_2)(L(x))=\alpha\cdot L_1(L(x))+\beta\cdot L_2(L(x))=\alpha\cdot\circ(L_1,L)(x)+\beta\cdot\circ(L_2,L)(x)=(\alpha\cdot\circ(L_1,L)+\beta\cdot\circ(L_2,L))(x)$ so that $\circ(\alpha\cdot L_1+\beta\cdot L_2)=\alpha\cdot\circ(L_1,L)+\beta\cdot\circ(L_2,L)$
2. Let $\alpha,\beta\in\mathbb{K},L\in L(Y,Z)$ and $L_1,L_2\in L(X,Y)$ then $\forall x\in X$ we have $\circ(L,\alpha\cdot L_1+\beta\cdot L_2)(x)=L((\alpha\cdot L_1+\beta\cdot L_2)(x))=L(\alpha\cdot L_1(x)+\beta\cdot L_2(x))=\alpha\cdot L(L_1(x))+\beta\cdot L(L_2(x))=\alpha\cdot\circ(L,L_1)(x)+\beta\cdot\circ(L,L_2)(x)=(\alpha\cdot\circ(L,L_1)+\beta\cdot\circ(L,L_2))(x)$ so that $\circ(L,\alpha\cdot L_1+\beta\cdot L_2)=\alpha\cdot\circ(L,L_1)+\beta\cdot\circ(L,L_2)$

Finally to prove continuity note $\|\|_1$, $\|\|_1$, $\|\|_3$ as the operator norm of $L(Y,Z)$, $L(X,Y),L(X,Z)$ then $\forall(L_1,L_2)\in L(Y,Z)\times L(X,Y)$ then using 12.183 we have that $\|L_1\circ L_2\|_3\leqslant\|L_1\|_1\cdot\|L_2\|_{2,}=1\cdot\prod_{i\in\{1,..,2\}}\|L_i\|_i$ proving that $\circ$ is continuous and that $\|\circ\|\leqslant 1$ □

Definition 12.197. *Let $n\in\mathbb{N}$ $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a finite family of normed spaces, $\langle Y,\|\|_Y\rangle$ a normed space then $L^n(X_1,\ldots,X_n;Y)$ is defined by*

1. *If $n=1$ then $L^n(X_1,\ldots,X_1;Y)=L(X_1,Y)$*
2. *If $n>1$ then $L^n(X_1,\ldots,X_n;Y)=L(X_1,L^{n-1}(X_2,\ldots,X_n;Y))$*

Example 12.198.

1. $L^1(X_1;Y)=L(X_1;Y)$
2. $L^2(X_1,X_2;Y)=L(X_1,L(X_2,Y))$

3. $L^3(X_1, ..., X_3) = L(X_1, L^2(X_2, X_3; Y)) = L(X_1, L(X_2, L^1(X_3, ..., X_3; Y))) = L(X_1, L(X_2, L(X_3, Y)))$
4. ...

Definition 12.199. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$, be normed vector spaces, $n \in \mathbb{N}$ then $L^n(X; Y)$ is defined as $L^n(X_1, ..., X_n; Y)$ where the family $\{X_i\}_{i \in \{1,...,n\}}$ is defined by $\forall i \in \{1, ..., n\}$ $X_i = X$*

1. $n = 1 \Rightarrow L^1(X; Y) = L(X, Y)$
2. $n > 1 \Rightarrow L^n(X; Y) = L(X, L^{n-1}(X; Y))$

Example 12.200.

1. $L^1(X; Y) = L(X, Y)$
2. $L^2(X: Y) = L(X, L^1(X; Y)) = L(X, L(X, Y))$
3. $L^3(X; Y) = L(X, L^2(X; Y)) = L(X, L(X, L(X, Y)))$
4. ...

Theorem 12.201. *Let $n \in \mathbb{N}$ $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1,...,n+1\}}$ a finite family of normed spaces, $\langle Y, \|\|_Y\rangle$ a normed space then $L^n(X_1, ..., X_n; L(X_{n+1}, Y)) = L^{n+1}(X_1, ..., X_{n+1}; Y)$*

Proof. We prove this by induction, so let $S = \{n \in \mathbb{N} | \text{for } \{\langle X_i, \|\|_i\rangle\}_{i \in \{1,...,n+1\}} \text{ we have } L^n(X_1, ..., X_n; L(X_{n+1}, Y)) = L^{n+1}(X_1, ..., X_{n+1}; Y)\}$ then:

1. If $n = 1$ then $L^1(X_1, ..., X_1; L(X_2, Y)) = L(X_1, L(X_2, Y)) = L^2(X_1, ..., X_2; Y)$ proving $1 \in S$
2. If $n \in S$ then $L^{n+1}(X_1, ..., X_{n+1}; L(X_{n+2}, Y)) = L(X_1, L^n(X_2, ..., X_{n+1}; L(X_{n+2}, Y))) \underset{n \in S}{=} L(X_1, L^{n+1}(X_2, ..., X_{n+2}; Y)) \underset{\text{def}}{=} L^{(n+1)+1}(X_1, ..., X_{n+2}; Y) \Rightarrow n + 1 \in S$

using mathematical induction proves the theorem. □

A easy corrolary of the above theorem is then

Corollary 12.202. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X\rangle$ a normed space, $\langle Y, \|\|_Y\rangle$ a normed vector space and $n \in \mathbb{N}$ then $L^n(X; L(X, Y)) = L^{n+1}(X; Y)$*

Definition 12.203. *Let $n \in \mathbb{N}$ $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1,...,n+1\}}$ a finite family of normed spaces, $\langle Y, \|\|_Y\rangle$ a normed vector space, $m \in \mathbb{N}$ with $m \leqslant n$ then if $L \in L^n(X_1, ..., X_n; Y)$ and $x = (x_1, ..., x_m) \in \prod_{i \in \{1,...,m\}} X_i$ we define $L(x_1: ...: x_m)$ recursively by*

1. $m = 1$ *then* $L(x_1: ...: x_1) = L(x_1)$
2. $m > 1$ *then* $L(x_1: ...: x_m) = (L(x_1: ...: x_{m-1}))(x_m)$

To show that the above is well defined we have the following theorem

Theorem 12.204. *Let $n\in\mathbb{N}$ $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a finite family of normed spaces, $\langle Y,\|\|_Y\rangle$ a normed vector space, $m\in\mathbb{N}$ with $m\leqslant n$ then if $L\in L^n(X_1,\ldots,X_n;Y)$ and $(x_1,\ldots,x_m)\in\prod_{i\in\{1,\ldots,m\}}$ then if $m<n$ we have that $L(x_1:\ldots:x_m)\in L^{n-m}(X_{m+1},\ldots,X_n;Y)$ and if $m=n$ then $L(x_1:\ldots:x_m)\in Y$*

Proof. We have two cases to consider

1. **$(m<n)$** we proceed by induction so let $S_n=\{m\in\mathbb{N}|\text{if } m<n \text{ then } L(x_1:\ldots:x_m)\in L^{n-m}(X_{m+1},\ldots,X_n;Y)\}$ then we must consider the following cases

    a. if $m=1$ then if $1<n$ we have $L(x_1:\ldots:x_m)=L(x_1:\ldots:x_1)=L(x_1)$ and as $L\in L^n(X_1,\ldots,X_n;Y)=L(X_1,L^{n-1}(X_2,\ldots,X_n:Y))$ we have that $L(x_1)\in L^{n-1}(X_2,\ldots,X_n;Y)$ proving that $1\in S_n$

    b. if $m\in S_n$ and $m+1<n\Rightarrow m<n$ then $L(x_1:\ldots:x_m)\in L^{n-m}(X_{m+1},\ldots,X_n;Y)\underset{m<n\Rightarrow n-m>1}{=}L(X_{m+1},L^{(n-m)-1}(X_{m+2},\ldots,X_n;Y))$ so that if $m+1<n$ then $L(x_1:\ldots:x_{m+1})=L(x_1:\ldots:x_m)(x_{m+1})\in L^{(n-m)-1}(X_{m+2},\ldots,X_n;Y)=L^{n-(m+1)}(X_{m+1},\ldots,X_n;Y)$ proving that $m+1\in S_n$

    By mathematical induction we have then that $S_n=\mathbb{N}$ so if $m<n\Rightarrow m\in S_n\wedge m<n$ so that $L(x_1:\ldots:x_m)\in L^{n-m}(X_{m+1},\ldots,X_n;Y)$

2. **$(m=n)$** Two cases must be considered

    a. **$(n=1)$** then $L(x_1:\ldots:x_1)=L(x_1)\in Y$ as $L\in L^1(X_1\ldots X_1;Y)=L(X_1,Y)$

    b. **$(n>1)$** then as $n-1<n$ we have by (1) $L(x_1:\ldots:x_{n-1})\in L^{n-(n-1)}(X_n,\ldots,X_n;Y)=L^1(X_n,\ldots,X_n;Y)=L(X_n,Y)$ so that $L(x_1,\ldots,x_n)=L(x_1,\ldots,x_{n-1})(x_n)\in Y$ □

The next theorem shows that $(x_1:\ldots:x_n)$ behaves the same as $(x_1,\ldots,x_n)$ functions.

Lemma 12.205. *Let $n\in\mathbb{N}$ $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a finite family of normed spaces, $\langle Y,\|\|_Y\rangle$ a normed vector space and $L,L'\in L^n(X_1,\ldots,X_n;Y)$ then if $\forall(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}X_i$ we have $L(x_1:\ldots:x_n)=L'(x_1:\ldots:x_n)$ then $L=L'$.*

Proof. We prove this by induction. So let $S_n=\left\{m\in\mathbb{N}|\text{if } m\leqslant n \text{ and } \forall(x_1,\ldots,x_m)\in\prod_{i\in\{1,\ldots,m\}}X_i \text{ we have } L(x_1:\ldots:x_m)=L'(x_1:\ldots:x_m) \text{ then } L=L'\right\}$ then we have

1. if $m=1$ and $\forall(x_1)\in\prod_{i\in\{1,\ldots,1\}}X_i$ we have $L(x_1)=L'(x_1)$ we have $\forall x\in X_1$ that $(x)\in\prod_{i\in\{1,\ldots,m\}}X_i$ and thus $L(x)=L(x')$ so that $L=L'$ and thus $1\in S_n$

2. if $m\in S_n$ then if $m+1\leqslant n$ and we have $\forall(x_1,\ldots,x_{m+1})\in\prod_{i\in\{1,\ldots,m+1\}}X_i$ we have $L(x_1:\ldots:x_{m+1})=L'(x_1:\ldots:x_{m+1})\Rightarrow L(x_1:\ldots:x_m)(x_{m+1})=L'(x_1:\ldots:x_m)(x_{m+1})\underset{x_{m+1}\,\text{is choosen arbitray}}{\Rightarrow}L(x_1:\ldots:x_m)=L'(x_1:\ldots:x_m)\underset{m\in S_n}{=}L=L'$ proving that $m+1\in S_n$ □

Finally we show that $(x_1)(x_2\colon \ldots.\colon x_n) = (x_1\colon \ldots\colon x_n)$

Theorem 12.206. *Let $n \in \mathbb{N} \setminus \{1\}$, $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1,\ldots,n\}}$ a finite family of normed spaces, $\langle Y, \|\|_Y\rangle$ a normed vector space and $L \in L^n(X_1, \ldots, X_n; Y)$ then $\forall (x_1, \ldots, x_n) \in \prod_{i \in \{1,\ldots,n\}} X_i$ we have $(L(x_1))(x_2\colon \ldots\colon x_n) = L(x_1\colon \ldots\colon x_n)$*

Proof. We proof this by induction so let $S = \{m \in \{2, \ldots\} | \text{if } m \leqslant n \text{ then } (L(x_1))(x_2\colon \ldots\colon x_m) = L(x_1\colon \ldots\colon x_m)\}$ then we have

1. If $m = 2$ then if $m \leqslant n$ we have $(L(x_1))(x_2\colon \ldots\colon x_m) = (L(x_1))(x_2\colon \ldots\colon x_2) = (L(x_1))(x_2) = L(x_1\colon \ldots\colon x_2)$ proving that $2 \in S$
2. If $m \in S$ then if $m + 1 \leqslant n$ we have $(L(x_1))(x_2\colon \ldots\colon x_{m+1}) = ((L(x_1))(x_2\colon \ldots\colon x_m))(x_{m+1}) \underset{m \in S}{=} (L(x_1\colon \ldots\colon x_m))(x_{m+1}) = L(x_1\colon \ldots\colon x_{m+1})$ proving that $m + 1 \in S$

So by induction we have that $S = \{2, \ldots\}$ so if $m = n > 1 \Rightarrow m \in \{2, \ldots\} = S \Rightarrow L(x_1)(x_2\colon \ldots\colon x_n) = L(x_1)(x_2\colon \ldots\colon x_m) = L(x_1\colon \ldots\colon x_m) = L(x_1\colon \ldots\colon x_n)$. □

Lemma 12.207. *Let $n \in \mathbb{N} \setminus \{1\}$, $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1,\ldots,n\}}$ a finite family of normed spaces, $\langle Y, \|\|_Y\rangle$ a normed vector space $\alpha, \beta \in \mathbb{K}$, $L_1, L_2 \in L^n(X_1, \ldots, X_n; Y)$, $x_1, \ldots, x_n \in X$ then $(\alpha \cdot L_1 + \beta \cdot L_2)(x_1\colon \ldots\colon x_n) = \alpha \cdot L_1(x_1\colon \ldots\colon x_n) + \beta \cdot L_2(x_1\colon \ldots\colon x_n)$*

Proof. We prove this by induction so let $S_n = \{m \in \mathbb{N} | \text{if } m < n \text{ then } (\alpha \cdot L_1 + \beta \cdot L_2)(x_1\colon \ldots\colon x_n) = \alpha \cdot L_1(x_1\colon \ldots\colon x_n) + \beta \cdot L_2(x_1\colon \ldots\colon x_n)\}$ then

1. if $m = 1$ $(\alpha \cdot L_1 + \beta \cdot L_2)(x_1\colon \ldots\colon x_1) = (\alpha \cdot L_1 + \beta \cdot L_2)(x_1) = \alpha \cdot L_1(x_1) + \beta \cdot L_2(x_1) = \alpha \cdot L_1(x_1\colon \ldots\colon x_1) + \beta \cdot L_2(x_1\colon \ldots\colon x_1)$ so $1 \in S_n$.
2. if $m \in S$ then $(\alpha \cdot L_1 + \beta \cdot L_2)(x_1\colon \ldots\colon x_{n+1}) = ((\alpha \cdot L_1 + \beta \cdot L_2)(x_1\colon \ldots\colon x_n))(x_{n+1}) \underset{m \in S}{=} (\alpha \cdot L_1(x_1\colon \ldots\colon x_n) + \beta \cdot L_2(x_1\colon \ldots\colon x_n))(x_{n+1}) = \alpha \cdot L_1(x_1\colon \ldots\colon x_n)(x_{n+1}) + \beta \cdot L_2(x_1\colon \ldots\colon x_n)(x_{n+1}) = \alpha \cdot L_1(x_1\colon \ldots\colon x_{n+1}) + \beta \cdot L_2(x_1\colon \ldots\colon x_{n+1})$

Using induction we have $S_n = \mathbb{N}$ proving the theorem. □

Lemma 12.208. *Let $n \in \mathbb{N} \setminus \{1\}$, $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1,\ldots,n\}}$ a finite family of normed spaces, $\langle Y, \|\|_Y\rangle$ a normed vector and $(x_1, \ldots, x_n) \in \prod_{i \in \{1,\ldots,\}} X_i$ then $\|L(x_1\colon \ldots\colon x_n)\|_Y \leqslant \|L\| \cdot \prod_{i \in \{1,\ldots,n\}} \|x_i\|_i$*

Proof. We must consider the following cases for n

1. **($\boldsymbol{n = 1}$)** Then as $L \in L^1(X_1; Y) = L(X_1, Y)$ we have $\forall (x_1) \in \prod_{i \in 1,\ldots,1} X_i$ that $\|L(x_1)\|_Y \leqslant \|L\| \cdot \|x_1\|_i = \|L\| \cdot \prod_{i=1} \|x_i\|_i$ proving the theorem in the case of $n = 1$
2. **($\boldsymbol{n > 1}$)** then if we note the norms in $L^{n-i}(X_{i+1}, \ldots, X_n; Y)$, $i \leqslant n$ by $\|\|_{(i)}$ and define $T_n = \left\{m \in \mathbb{N} | \text{if } m \leqslant n - 1 \text{ then } \forall (x_1, \ldots, x_m) \in \prod_{i \in \{1,\ldots,m\}} X_m \text{ we have } \|L(x_1\colon \ldots\colon x_m)\|_{(m)} \leqslant \|L\| \cdot \prod_{i \in \{1,\ldots,m\}} \|x_i\|_i\right\}$ then we have:
    a. If $m = 1$ then as $n \geqslant 2$ we have $m < n - 1$ and $\|L(x_1)\|_{(1)} \leqslant \|L\| \cdot \|x_1\|_1 = \|L\| \cdot \prod_{i \in \{1,\ldots,1\}} \|x_i\|_i \Rightarrow 1 \in T_n$

b. If $m \in T_n$ then if $m+1 \leqslant n-1$ and $(x_1, \ldots, x_{m+1}) \in \prod_{i \in \{1, \ldots, m+1\}} X_i$ then $\|L(x_1: \ldots: x_{m+1})\|_{(m+1)} = \|L(x_1: \ldots: x_m)(x_{m+1})\|_{(m+1)} \leqslant \|L(x_1: \ldots: x_m)\|_{(m)} \cdot \|x_{m+1}\|_{m+1} \leqslant_{m \in T_n} \left(\|L\| \cdot \prod_{i \in \{1, \ldots, m\}} \|x_i\|_i\right) = \|L\| \cdot \prod_{i \in \{1, \ldots, m+1\}} \|x_i\|_i$ proving that $m+1 \in T_n$

Using mathematical induction we have that $T_n = \mathbb{N}$ so for $n-1 \in \mathbb{N} = T_n$ we have as $n-1 \leqslant n-1$ that $\|L(x_1: \ldots: x_{n-1})\|_{(n-1)} \leqslant \|L\| \cdot \prod_{i \in \{1, \ldots, n-1\}} \|x_i\|_i$ so that $\|L(x_1: \ldots: x_n)\|_Y = \|L(x_1: \ldots: x_{n-1})(x_n)\| \leqslant \|L(x_1: \ldots: x_{n-1})\|_{(n-1)} \cdot \|x_n\|_n \leqslant \left(\|L\| \cdot \prod_{i \in \{1, \ldots, n-1\}} \|x_i\|_i\right) \cdot \|x_n\|_n = \|L\| \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i$ proving the theorem for $n > 1$ □

Theorem 12.209. *Let $\langle \mathbb{K}, \| \| \rangle$, $\langle Y, \| \|_Y \rangle$ be normed spaces (where $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$ with the absolute norm), $n \in \mathbb{N}$ and $L \in L^n(\mathbb{K}, Y)$ then $\|L\|_{L^n(\mathbb{K};Y)} \leqslant \|L(1: \ldots: 1)\|_Y$*

Proof. We prove this by induction on n , denote $\| \|_i = \| \|_{L^i(\mathbb{K};Y)}$ the operator norm in $L^i(\mathbb{K};Y)$ and let $S = \{n \in \mathbb{N} | L \in L^n(\mathbb{K}, Y) \text{ then } \|L\|_i \leqslant \|L(1: \ldots: 1)\|_Y\}$ then we have:

1. If $n = 1$ then $L \in L^1(\mathbb{K};Y) = L(\mathbb{K}, Y)$ then if $|x_1| = 1$ we have as $x_1 = x_1 \cdot 1$ that $\|L(x_1)\|_Y = \|x_1 \cdot L(1)\|_Y = |x_1| \cdot \|L(1)\|_Y = \|L(1)\| \leqslant \|L(1)\|_Y$ and as $\|L\|_1 \underset{12.178}{=} \inf(\{k \in \mathbb{R}_+ | \forall x \in \mathbb{K} \vdash |x| = 1 \text{ we have } \|L(x)\|_Y \leqslant k\})$ that $\|L\|_1 \leqslant \|L(1)\|_Y$ proving that $1 \in S$
2. If $n \in S$ then as $L^{n+1}(\mathbb{K};Y) = L(\mathbb{K}, L^n(\mathbb{K};Y))$ we have if $L \in L^{n+1}(\mathbb{K};Y)$ that $L(1) \in L^n(\mathbb{K};Y)$ so that as $n \in S$ we have that

$$\|L(1)\|_n \leqslant \left\| L(1)\underbrace{(1: \ldots: 1)}_{n} \right\|_Y = \left\| L\underbrace{(1: \ldots: 1)}_{n+1} \right\|_Y \tag{12.22}$$

now if $x \in \mathbb{K}$ is such that $|x| = 1$ then $\|L(x)\|_Y = \|L(x \cdot 1)\|_Y = \|x \cdot L(1)\|_Y = |x| \cdot \|L(1)\|_Y = \|L(1)\|_Y$ which as $\|L\|_{n+1} = \inf(\{k \in \mathbb{R}_+ | \forall x \in \mathbb{K} \vdash |x| = 1 \text{ we have } \|L(x)\|_Y \leqslant k\})$ we have that $\|L\|_{n+1} \leqslant \|L(1)\|_Y$ proving using 12.23 that $\|L\|_{n+1} \leqslant \left\| L\underbrace{(1: \ldots: 1)}_{n+1} \right\|_Y$ and thus that $n+1 \in S$

Mathematical induction then proves that $S = \mathbb{N}$ proving the theorem. □

Definition 12.210. *Let $n \in \mathbb{N}$, $\{\langle X_i, \| \|_i \rangle\}_{i \in \{1, \ldots, n\}}$ a finite family of normed spaces, $\langle Y, \| \|_Y \rangle$ a normed vector then given $L \in L^n(X_1, \ldots, X_n; Y)$ we define $\mathcal{P}_n(L)$: $\prod_{i \in \{1, \ldots, n\}} X_i \rightarrow Y$ by $\mathcal{P}_n(L)(x_1, \ldots, x_n) = L(x_1: \ldots: x_n)$. Note that if $n = 1$ then $L^1(X_1, \ldots, X_1; Y) = L(X_1, Y) = L(X_1 \ldots, X_!; Y)$ and then $\mathcal{P}_n(L) = 1_{L(X_1, Y)}$.*

Theorem 12.211. *Let $n \in \mathbb{N}$, $\{\langle X_i, \| \|_i \rangle\}_{i \in \{1, \ldots, n+1\}}$ a finite family of normed spaces, $\langle Y, \| \|_Y \rangle$ a normed vector then given $L \in L^{n+1}(X_1, \ldots, X_{n+1}; Y) \underset{12.201}{=} L^n(X_1, \ldots, X_n; L(X_{n+1}, Y))$ and $(x_1, \ldots, x_{n+1}) \in \Pi_{i \in \{1, \ldots, n+1\}} X_i$ then $\mathcal{P}_{n+1}(L)(x_1, \ldots, x_{n+1}) = \mathcal{P}_n(L)(x_1, \ldots, x_n)(x_{n+1})$*

Proof. $\mathcal{P}_{n+1}(L)(x_1, \ldots, x_{n+1}) = L(x_1: \ldots: x_{n+1}) = L(x_1: \ldots: x_n)(x_{n+1}) = \mathcal{P}_n(L)(x_{n+1})$ □

Theorem 12.212. *Let $n\in\mathbb{N}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n+1\}}$ a finite family of normed spaces, $\langle Y,\|\|_Y\rangle$ a normed vector and $L\in L^n(X_1,\ldots,X_n;Y)$ then $\mathcal{P}_n(L)$ is multilinear and continuous (so that $\mathcal{P}_n(L)\in L(X_1,\ldots,X_n;Y)$ where $\|\mathcal{P}_n(L)\|_0\leqslant\|L\|$ (where $\|\|$ is the operator norm of $L^n(X_1,\ldots,X_n;Y)=\left\{\begin{array}{l} L(X_1,Y)\text{ if } n=1\\ L(X_1,L^{n-1}(X_2,\ldots,X_n;Y))\text{ if } n>1\end{array}\right.$) and $\|\|_0$ the operator norm on $L(X_1,\ldots,X_n;Y)$.*

Proof. We divide the proof in different phases

1. **(multilinearity)** Let $S=\{m\in\mathbb{N}|\text{if } m\leqslant n \text{ then if } i\in\{1,\ldots,m\} \text{ we have } \forall\alpha,\beta\in\mathbb{K},\forall x,y\in X_i,\forall z\in\prod_{j\in\{1,\ldots,n\}\setminus\{i\}}X_j \text{ that } \mathcal{P}_m(L)(z_1,\ldots,z_{i-1},\alpha\cdot x+\beta\cdot y,z_{i+1},\ldots,z_m)=\alpha\cdot\mathcal{P}_m(L)(z_1,\ldots,z_{i-1},x,z_{i+1},\ldots,z_m)+\beta\cdot\mathcal{P}_m(z_1,\ldots,z_{i-1},y,z_{i+1},\ldots,z_m)\}$ then we have
   a. If $m=1$ (so $m\leqslant n$) then as $\mathcal{P}_1(L)=L$ and as L is linear we have $1\in S$
   b. If $m\in S$ then if $m+1\leqslant n$ we have two cases to consider for $i\in\{1,\ldots,m+1\}$
      i. **($i\leqslant m$)** then as $m\leqslant n$ we have $\mathcal{P}_{m+1}(L)(z_1,\ldots,z_{i-1},\alpha\cdot x+\beta\cdot y,z_{i+1},\ldots,z_{m+1})\underset{12.211}{=}\mathcal{P}_m(L)(z_1,\ldots,z_{i-1},\alpha\cdot x+\beta\cdot y,z_{i+1},\ldots,z_m)(z_{m+1})\underset{m\in S\wedge m<n}{=}(\alpha\cdot\mathcal{P}_m(L)(z_1,\ldots,z_{i-1},x,z_{i+1},\ldots,z_m)+\beta\cdot\mathcal{P}_m(z_1,\ldots,z_{i-1},y,z_{i+1},\ldots,z_m))(z_{m+1})=\alpha\cdot\mathcal{P}_m(L)(z_1,\ldots,z_{i-1},x,z_{i+1},\ldots,z_m)(z_{m+1})+\beta\cdot\mathcal{P}_m(z_1,\ldots,z_{i-1},y,z_{i+1},\ldots,z_m)(z_{m+1})=\alpha\cdot\mathcal{P}_{m+1}(L)(z_1,\ldots,z_{i-1},x,z_{i+1},\ldots,z_{m+1})+\beta\cdot\mathcal{P}_{m+1}(z_1,\ldots,z_{i-1},y,z_{i+1},\ldots,z_{m+1})$
      ii. **($i=m+1$)** then $\mathcal{P}_{m+1}(L)(z_1,\ldots,z_{i-1},\alpha\cdot x+\beta\cdot y,z_{i+1},\ldots,z_n)=\mathcal{P}_{m+1}(L)(z_1,\ldots,z_m,\alpha\cdot x+\beta\cdot y)\underset{12.211}{=}\mathcal{P}_m(L)(z_1,\ldots,z_m)(\alpha\cdot x+\beta\cdot y)=L(z_1:\ldots z_m)(\alpha\cdot x+\beta\cdot y)\underset{L(z_1:\ldots:z_m)\in L^{n-m}(X_{m+1},\ldots,X_n;Y)}{=}\alpha\cdot L(z_1:\ldots:z_m)(x)+\beta\cdot L(z_1:\ldots:z_m)(y)=\alpha\cdot L(z_1:\ldots:z_m:x)+\beta\cdot L(z_1:\ldots:z_m:y)=\alpha\cdot\mathcal{P}_{m+1}(L)(z_1,\ldots,z_{i-1},x,z_{i+1},\ldots,z_{m+1})+\beta\cdot\mathcal{P}_{m+1}(z_1,\ldots,z_{i-1},y,z_{i+1},\ldots,z_{m+1})$

   (i) and (2) proves then that $m+1\in S$

   By mathematical induction we have then that $S=\mathbb{N}$ so as $n\in\mathbb{N}=S$ and $n\leqslant n$ we have proved multilinearity.

2. **(continuity)** If $(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}X_i$ then we have by 12.208 that $\|\mathcal{P}(L)(x_1,\ldots,x_n)\|_Y=\|L(x_1:\ldots:x_n)\|_Y\leqslant\|L\|\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i$ proving that as $\mathcal{P}(L)$ is multilinear that $\mathcal{P}(L)$ is continuous and that $\|\mathcal{P}(L)\|\leqslant\|L\|$ (see 12.191) □

Theorem 12.213. *Let $n\in\mathbb{N}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n+1\}}$ a finite family of normed spaces, $\langle Y,\|\|_Y\rangle$ a normed vector and $L\in L^n(X_1,\ldots,X_n;Y)$ then $\mathcal{P}_n: L^n(X_1,\ldots,X_n;Y)\to L(X_1,\ldots,X_n;Y)$ defined by $\mathcal{P}_n(L)(x_1,\ldots,x_n)=L(x_1:\ldots:x_n)$ (see 12.210) is a norm preserving isomorphism*

Proof. *The proof is done in different steps*

1. **(injectivity)** *If $\mathcal{P}_n(L)=\mathcal{P}_n(L')$ then $\forall x_1,\ldots,x_n\in X$ we have $L(x_1:\ldots:x_n)=\mathcal{P}_n(L)=\mathcal{P}_n(L')=L'(x_1:\ldots:x_n)\underset{12.205}{\Rightarrow}L=L'$ proving injectivity.*

2. **(surjectivity)** *Let* $S=\{n\in\mathbb{N}|\text{if } \{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n+1\}}$ is *a* family of normed spaces and $\langle Y,\|\|_Y\rangle$ is *a* normed space then $\mathcal{P}_n\colon L^n(X_1,\ldots,X_n;Y)\to L(X_1,\ldots,X_n;Y)$ is *a* surjection$\}$ *then we have*

   a. *If* $n=1$ *then* $L^1(X_1;Y)=L(X_1,Y)=L(X_1,\ldots,X_1;Y)$ *and* $\mathcal{P}_1=1_{L(X_!,\ldots,X_1;Y)}$ *which is clearly surjective so that* $1\in S$

   b. *If* $n\in S$ *then let* $L\in L(X_1,\ldots,X_{n+1};Y)$ *then* $\forall(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}X_i$ *define then* $L_1\colon\prod_{i\in\{1,\ldots,n\}}X_i\to L(X_{n+1},Y)$ *by* $(x_1,\ldots,x_n)\to L_1(x_1,\ldots,x_n)$ *where* $L_1(x_1,\ldots,x_n)(x)=L(x_1,\ldots,x_n,x)$. *Of course we must prove that* $L_1(x_1,\ldots,x_n)\in L(X_{n+1},Y)$

      i. **(linearity)** *If* $\alpha,\beta\in\mathbb{K}$ *and* $x,y\in X_{n+1}$ *then* $L_1(x_1,\ldots,x_n)(\alpha\cdot x+\beta\cdot y)=L(x_1,\ldots,x_n,\alpha\cdot x+\beta\cdot y)\underset{L\,\text{is multilinear}}{=}\alpha\cdot L(x_1,\ldots,x_n,x)+\beta\cdot L(x_1,\ldots,x_n,y)=\alpha\cdot L_1(x_1,\ldots,x_n)(x)+\beta\cdot L_1(x_1,\ldots,x_n)(y)$ *proving that* $L_1(x_1,\ldots,x_n)$ *is linear.*

      ii. **(continuity)** *Let* $x\in X_{n+1}$ *then* $\|L_1(x_1,\ldots,x_n)(x_{n+1})\|_Y=\|L(x_1,\ldots,x_n,x)\|_Y\leqslant\left(\|L\|\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i\right)\cdot\|x\|$ *proving continuity.*

   Further the mapping L_1 *itself is multilinear and continuous*

      i. **(multilinearity)** *If* $i\in\{1,\ldots,n\}$, $x,y\in X_i$, $\alpha,\beta\in\mathbb{K}$, $(z_1,\ldots,z_n)\in\prod_{j\in\{1,\ldots n\}\setminus\{i\}}X_j$ *and* $w\in X_{n+1}$ *then* $L_1(z_1,\ldots,z_{i-1},\alpha\cdot x+\beta\cdot y,z_{i+1},\ldots,z_n)(w)=L(z_1,\ldots,x_{i-1},\alpha\cdot x+\beta\cdot y,z_{i+1},\ldots,z_n,w)\underset{L\,\text{is multilinear}}{=}\alpha\cdot L(z_1,\ldots,z_{i-1},x,z_{i+1},\ldots,z_n,w)+\beta\cdot L(z_1,\ldots,z_{i-1},y,z_{i+1},\ldots,z_n,w)=\alpha\cdot L_1(z_1,\ldots,z_{i-1},x,z_{i+1},\ldots.,z_n)(w)+\beta\cdot L_1(z_1,\ldots,z_{i-1},y,z_{i+1},\ldots,z_n)(w)$ *proving as* w *is choosen arbitrary we have that* $L_1(z_1,\ldots,z_{i-1},\alpha\cdot x+\beta\cdot y,z_{i+1},\ldots,z_n)=\alpha\cdot L_1(z_1,\ldots,z_{i-1},x,z_{i+1},\ldots,z_n)+\beta\cdot L_1(z_1,\ldots,z_{i-1},y,z_{i+1},\ldots,z_n)$

      ii. **(continuity)** $\forall(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}X_i$, $\forall x\in X_{n+1}$ *we have that* $\|L_1(x_1,\ldots,x_n)(x)\|_Y=\|L(x_1,\ldots,x_n,x)\|_Y\leqslant\left(\|L\|\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i\right)\cdot\|x\|_{n+1}$ *so that* $\|L_1(x_1,\ldots,x_n)\|_{L(X_{n+1},Y)}\leqslant\|L\|\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|$ *and thus proving that* L_1 *is continuous.*

   So we conclude that $L_1\in L(X_1,\ldots,X_n;L(X_{n+1},Y))$ *then as* $n\in S$ *there exists a* $L_2\in L^n(X_1,\ldots,X_n;L(X_{n+1},Y))\underset{12.201}{=}$ *such that* $\mathcal{P}_n(L_2)=L_1$. *Then* $\forall(x_1,\ldots,x_{n+1})\in\prod_{i\in\{1,\ldots,n+1\}}X_i$ *we have* $L(x_1,\ldots,x_{n+1})=L_1(x_1,\ldots,x_n)(x_{n+1})=(\mathcal{P}_n(L_2)(x_1,\ldots,x_n))(x_{n+1})=L_2(x_1\colon\ldots;x_n)(x_{n+1})=L_2(x_1\colon\ldots\colon x_{n+1})=\mathcal{P}_{n+1}(L_2)(x_1,\ldots,x_{n+1})$ *proving that* $L=\mathcal{P}_{n+1}(L_2)$ *and thus that* $\mathcal{P}_{n+1}$ *is surjective.*

3. **($\boldsymbol{\mathcal{P}_n}$ is linear)** *So let* $\alpha,\beta\in\mathbb{K}$ *and* $L_1,L_2\in L^n(X_1,\ldots,X_n;Y)$ *then* $\forall(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}X_i$ *we have* $(\mathcal{P}_n(\alpha\cdot L_1+\beta\cdot L_2))(x_1,\ldots,x_n)=(\alpha\cdot L_1+\beta\cdot L_2)(x_1\colon\ldots\colon x_n)\underset{12.207}{=}\alpha\,L_1(x_1\colon\ldots\colon x_n)+\beta\cdot L_2(x_1\colon\ldots\colon x_n)=\alpha\cdot\mathcal{P}_n(L_1)(x_1,\ldots,x_n)+\beta\cdot\mathcal{P}_n(L_2)(x_1,\ldots,x_n)=(\alpha\cdot\mathcal{P}_n(L_1)+\beta\cdot\mathcal{P}_n(L_2))(x_1,\ldots,x_n)$ *proving that* $\mathcal{P}_n(\alpha\cdot L_1+\beta\cdot L_2)=\alpha\cdot\mathcal{P}_n(L_1)+\beta\cdot\mathcal{P}_n(L_2)$

4. **(Norm preserving)** *By the previous theorem we have already that if* $L \in L^n(X_1, \ldots, X_n; Y)$ *that* $\|\mathcal{P}_n(L)\|_{L(X_1, \ldots, X_n; Y)} \leqslant \|L\|$. *We prove now by induction that* $\|L\| \leqslant \|\mathcal{P}_n(L)\|_{L(X_1, \ldots, X_n; Y)} \leqslant \|L\|$. *Define given* $i \in \{0, \ldots, n\}$ *that* $\|\|_{(i)} = \begin{cases} \|\|_Y \text{ if } i = 0 \\ \|\|_{L^i(X_{n-(i-1)}, \ldots, X_n; Y)} \end{cases}$ *and take then* $S_n = \{m \in \mathbb{N}_0 \,|\text{if } m < n \text{ then } \forall (x_1, \ldots, x_{n-m}) \in \prod_{i \in \{1, \ldots, n-m\}} X_i \text{ with } \forall i \in \{1, \ldots, n-m\} \vdash \|x_i\|_i = 1 \text{ we have } \|L(x_1 \colon \ldots \colon x_{n-m})\|_{(m)} \leqslant \|\mathcal{P}_n(L)\|_{L(X_1, \ldots, X_n; Y)}\}$ *then we have :*

   a. *If* $m = 0$ *then* $m < n$ *and* $n - m = n$ *so if* $(x_1, \ldots, x_n) \in \prod_{i \in \{1, \ldots, n-0\}} X_i$ *with* $\forall i \in \{1, \ldots, n\}$ $\|x_i\| = 1$ *we have* $\|L(x_1 \colon \ldots \colon x_n)\|_0 = \|L(x_1 \colon \ldots \colon x_n)\|_Y = \|\mathcal{P}_n(L)(x_1, \ldots, x_n)\|_Y \leqslant \|\mathcal{P}_n(L)\|_{L(X_1, \ldots, X_n; Y)} \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i = \|\mathcal{P}_n(L)\|_{L(X_1, \ldots, X_n; Y)}$ *proving that* $0 \in S_n$

   b. *If* $m \in S_n$ *then if* $m + 1 < n \Rightarrow 0 < n - (m+1)$ *we have* $\forall (x_1, \ldots, x_{n-(m+1)}) \in \prod_{i \in \{1, \ldots, n-(m+1)\}} X_i$ *with* $\|x_i\|_i = 1$ *if* $i \in \{1, \ldots, n - (m+1)\}$ *we have using 12.181 that* $\|L(x_1 \colon \ldots \colon x_{n-(m+1)})\|_{(m+1)} = \sup(\{\|L(x_1 \colon \ldots \colon x_{n-(m+1)})(x)\|_{(m)} | x \in X_{n-m} \text{ with } \|x\|_{n-m} = 1\}) = \sup(\{\|L(x_1 \colon \ldots \colon x_{n-(m+1)} \colon x)\|_{(m)} | x \in X_{n-m} \text{ with } \|x\|_{n-m} = 1\}) \leqslant \|\mathcal{P}_n(L)\|_{L(X_1, \ldots, X_n; Y)}$ *[as* $m \in S$ *we have that as* $\forall z = (x_1, \ldots, x_{n-(m+1)}, x) \in \prod_{i \in \{1, \ldots, n-m\}} X_i$ *that* $\|z_i\|_i = 1$ *that* $\|L(x_1, \ldots, x_{n-(m+1)}, x)\|_{(m)} = \|L(z_1 \colon \ldots \colon z_{n-m})\|_{(m)} \leqslant \|\mathcal{P}_n(L)\|_{L(X_1, \ldots, X_n \colon Y)}$*] so that* $m + 1 \in S_n$.

   By induction we have then that $S_n = \mathbb{N}_0 \Rightarrow n - 1 \in S$ *so as* $n - 1 < n$ *we have that* $\forall x \in X_1 | \vdash \|x\|_1 = 1 \Rightarrow \forall (x) \in \prod_{i \in \{1, \ldots, n-(n-1)\}} X_i$ *with* $\|x\|_1 = 1$ *that* $\|L(x_1)\|_{(n-1)} = \|L(x_1 \colon \ldots \colon x_1)\|_{(n-1)} \leqslant \|\mathcal{P}_n(L)\|$. *As* $\|L\| = \sup(\|L(x)\|_{(n-1)} | x \in X_1 \vdash \|x\|_1 = 1)$ *we have that* $\|L\| \leqslant \|\mathcal{P}_n(L)\|$. *This together with* $\|\mathcal{P}_n(L)\|_{L(X_1, \ldots, X_n; Y)} \leqslant \|L\|$ *proves that* $\|\mathcal{P}_n(L)\|_{L(X_1, \ldots, X_n; Y)} = \|L\|$. □

The above theorem lets us to identify $L^n(X; Y)$ with $L(X^n; Y)$ which will be used for higher order differentation.

Definition 12.214. *Let* $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$, $\langle Z, \|\|_Z \rangle$ *be normed spaces,* $n \in \mathbb{N}$ *and* $L \in L(Y, Z)$ *then given* $K \in L^n(X; Y)$ *we have* $\mathcal{P}_{n,Y}(K) \in L(X^n; Y) \Rightarrow L \circ \mathcal{P}_{n,Y}(K) \in L(X^n; Z) \Rightarrow \mathcal{P}_{n,Z}^{-1}(L \circ \mathcal{P}_{n,Y}(K)) \in L^n(X; Z)$ *we define then* $L \bullet_n K = \mathcal{P}_{n,Z}^{-1}(L \circ \mathcal{P}_{n,Y}(K))$. *If* $n = 1$ *then* $L \bullet_1 K = L \circ K$

Theorem 12.215. *Let* $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$, $\langle Z, \|\|_Z \rangle$ *be normed spaces and* $L \in L(Y, Z)$ *then given* $n \in \mathbb{N}$ *then* $K \in L^n(X; Y)$ *we have if* $x_1, \ldots, x_n \in X$ *that* $(L \bullet_n K)(x_1 \colon \ldots \colon x_n) = L(K(x_1 \colon \ldots \colon x_n))$

Proof. $(L \bullet_n K)(x_1 \colon \ldots \colon x_n) = \mathcal{P}_{n,Z}^{-1}(L \circ \mathcal{P}_{n,Y}(K))(x_1 \colon \ldots \colon x_n) \underset{\text{definition of } \mathcal{P}_{n,Z}}{=} (L \circ \mathcal{P}_{n,Y}(K))(x_1, \ldots, x_n) = L(\mathcal{P}_{n,Y}(K)(x_1, \ldots, x_n)) = L(K(x_1 \colon \ldots \colon x_n))$ □

Theorem 12.216. *Let* $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$, $\langle Z, \|\|_Z \rangle$ *be normed spaces and* $L \in L(Y, Z)$ *then given* $\mathrm{ev}_{L,n} \colon L^n(X; Y) \to L^n(X; Z)$ *defined by* $K \to L \bullet_n K$ *is a linear continuous function.*

Proof.

1. **(Linearity)** Let $K_1, K_2 \in L^n(X; Y)$ and $\alpha, \beta \in \mathbb{K}$ then we have that $\mathrm{ev}_L(\alpha \cdot K_1 + \beta \cdot K_2) = L \bullet_n (\alpha \cdot K_1 + \beta \cdot K_2) = \mathcal{P}_{n,Z}^{-1}(L \circ \mathcal{P}_{n,Y}(\alpha \cdot K_1 + \beta \cdot K_2)) \underset{\mathcal{P}_n \text{ is linear (see 12.213)}}{=} \mathcal{P}_{n,Z}^{-1}(L \circ (\alpha \cdot \mathcal{P}_{n,Y}(K_1) + \beta \cdot \mathcal{P}_{n,Y}(K_2))) \underset{L \text{ is linear}}{=} \mathcal{P}_{n,Z}^{-1}(\alpha \cdot (L \circ \mathcal{P}_{n,Y}(K_1)) + \beta \cdot (L \circ \mathcal{P}_{n,Y}(K_2))) \underset{\mathcal{P}_{n,Z} \text{ is a linear isomorphism} + 10.181}{=} \alpha \cdot \mathcal{P}_{n,Z}^{-1}(L \circ \mathcal{P}_{n,Y}(K_1)) + \beta \cdot \mathcal{P}_{n,Z}^{-1}(L \circ \mathcal{P}_{n,Y}(K_2)) = \alpha \cdot \mathrm{ev}_L(K_1) + \beta \cdot \mathrm{ev}_L(K_2)$
2. **(continuity)** Let $K \in L^n(X, Y)$ then $\|\mathrm{ev}_L(K)\|_{L^n(X;Y)} = \|\mathcal{P}_{n,Z}^{-1}(L \circ \mathcal{P}_{n,Y}(K))\|_{L^n(X;Y)} \underset{\mathcal{P}_{n,Z} \text{ is norm preserving}}{=} \|\mathcal{P}_{n,Z}(\mathcal{P}_{n,Z}^{-1}(L \circ \mathcal{P}_{n,Y}(K)))\|_{L(X^n;Z)} = \|L \circ \mathcal{P}_{n,Y}(K)\|_{L(X^n;Z)} \leqslant_{12.193} \|L\|_{L(Y,Z)} \cdot \|\mathcal{P}_{n,Y}(K)\| \underset{\mathcal{P}_{n,Y} \text{ is norm preserving}}{=} \|L\|_{L(X,Z)} \cdot \|K\|_{L^n(X;Y)}$ proving continuity. □

The next theorem shows how $\mathrm{ev}_{L,n}$ can be used to go one level higher

Theorem 12.217. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$, $\langle Z, \|\|_Z\rangle$ be normed spaces and $L \in L(Y, Z)$ then given $n \in \mathbb{N}$ we have if $K \in L^{n+1}(X;Y)$ that $\mathrm{ev}_{L,n} \circ K = L \bullet_{n+1} K$*

Proof. Let $x_1, \ldots, x_{n+1} \in X$ then $K(x_1) \in L(X, L^n(X;Y))$ so that $(\mathrm{ev}_{L,n} \circ K)(x_1) = \mathrm{ev}_{L,n}(K(x_1)) \underset{\text{definition of } \mathrm{ev}_{L,n}}{=} L \bullet_n K(x_1)$ so that $(\mathrm{ev}_{L,n} \circ K)(x_1: \ldots: x_{n+1}) = (\mathrm{ev}_{L,n} \circ K)(x_1)(x_2: \ldots: x_n) = (L \bullet_n K(x_1)) \underset{12.215}{=} L(K(x_1)(x_2: \ldots: x_{n+1})) = L(K(x_1: \ldots: x_{n+1})) \underset{12.215}{=} (L \bullet_{n+1} K)(x_1: \ldots: x_{n+1})$ so that $(\mathrm{ev}_{L,n} \circ K)(x_1: \ldots: x_{n+1}) = (L \bullet_{n+1} K)(x_1: \ldots: x_{n+1})$ so as $x_1, \ldots, x_n \in X$ is choosen arbitrary we have by 12.205 that $\mathrm{ev}_{L,n} \circ K = L \bullet_{n+1} K$. □

12.8 Separation

Definition 12.218. *A topological space $\langle X, \mathcal{T}\rangle$ is **Hausdorff** if $\forall x, y \in X$ with $x \neq y$ there exists a $U, V \in \mathcal{T}$ with $x \in U \wedge y \in V \wedge U \bigcap V = \emptyset$*

Example 12.219. A metric space $\langle X, d\rangle$ is Hausdorff in the metric topology

Proof. If $x, y \in X$ with $x \neq y$ then $\varepsilon = d(x, y) > 0$ define then $x \in B_d\left(x, \frac{\varepsilon}{2}\right)$ and $y \in B_d\left(y, \frac{\varepsilon}{2}\right)$. If now $z \in B_d\left(x, \frac{\varepsilon}{2}\right) \bigcap B_d\left(y, \frac{\varepsilon}{2}\right) \Rightarrow \varepsilon = d(x, y) \leqslant d(x, z) + d(y, z) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \Rightarrow \varepsilon < \varepsilon$ a contradiction so that $B_d\left(x, \frac{\varepsilon}{2}\right) \bigcap B_d\left(y, \frac{\varepsilon}{2}\right)$ □

Theorem 12.220. *Every finite subset of a Hausdorff space $\langle X, \mathcal{T}\rangle$ is closed.*

Proof. First note that the finite set $\emptyset$ is closed as $X \setminus \emptyset = X$ is open. Second we prove that if $x \in X$ then $\{x\}$ is closed or equivalent that $X \setminus \{x\}$ is open, so let $y \in X \setminus \{x\} \Rightarrow y \neq x \Rightarrow \exists U, V$ open with $x \in U, y \in V$ and $U \bigcap V = \emptyset \Rightarrow x \notin V \Rightarrow y \in V \subseteq X \setminus \{x\} \Rightarrow X \setminus \{x\}$ is open. If now F is a finite non empty set then there exists a bijection $b: \{0, \ldots, n\} \rightarrow F$ so that $F = \bigcup_{i \in \{1, \ldots, n\}} \{b(i)\} \Rightarrow F$ is closed as a finite set of closed sets is closed. □

Definition 12.221. *A topological space $\langle X, \mathcal{T}\rangle$ is* ***regular*** *if for every closed set A and $\forall x \in X$ with $x \notin A$ there exists disjoint open sets $U, V \in \mathcal{T}$ with $x \in U, A \subseteq V$ and $U \bigcap V = \emptyset$*

Theorem 12.222. *A metric space $\langle X, d\rangle$ is regular*

Proof. Let A be a closed set and $x \notin A$ then as $x \in X \setminus A$ (which is open) there exists a U open with $x \in U \subseteq X \setminus A$ so that $x \in U$ and $U \bigcap A = \emptyset$. As $x \in U$ open there exist a $\delta \in \mathbb{R}_+$ such that $x \in B_d(x, \delta) \subseteq U \Rightarrow B_d(x, \delta) \bigcap A = \emptyset$. Given $a \in A$ assume that $z \in B_d\left(x, \frac{\delta}{2}\right) \bigcap B_d\left(a, \frac{\delta}{2}\right) \Rightarrow d(x, z) < \frac{\delta}{2} \wedge d(z, a) < \frac{\delta}{2} \Rightarrow d(x, a) \leqslant d(x, z) + d(z, a) < \frac{\delta}{2} + \frac{\delta}{2} = \delta \Rightarrow a \in B_d(x, \delta) \Rightarrow a \in B_d(x, \delta) \bigcap A = \emptyset$ a contradiction so we must have that $B_d\left(x, \frac{\delta}{2}\right) \bigcap B_d\left(a, \frac{\delta}{2}\right) = \emptyset$ and thus as $A \subseteq \bigcup_{a \in A} B_d\left(a, \frac{\delta}{2}\right) = V$ a open set with $V \bigcap B_d\left(x, \frac{\delta}{2}\right) = \bigcup_{a \in A}\left(B_d\left(a, \frac{\delta}{2}\right) \bigcap B_d\left(x, \frac{\delta}{2}\right)\right) = \bigcup_{a \in A} \emptyset = \emptyset$ so $A \subseteq V$ a open set that does not intersect the open set $B_d\left(x, \frac{\delta}{2}\right)$ containing x. □

Theorem 12.223. *Let $\langle X, \mathcal{T}\rangle$ be a regular topological space and U a open set with $x \in U$ then there exists a open set V with $x \in V \wedge \bar{V} \subseteq U$*

Proof. Take $x \in U$ then as $x \notin X \setminus U$ a closed set we have by regularity a V, W open with $x \in V$, $X \setminus U \subseteq W$ and $V \bigcap W = \emptyset \Rightarrow V \subseteq X \setminus W \underset{X \setminus W \text{ is closed}}{\Rightarrow} \bar{V} \subseteq X \setminus W \underset{X \setminus U \subseteq W \Rightarrow X \setminus W \subseteq X \setminus (X \setminus U) = U}{\Rightarrow} \bar{V} \subseteq U$ and $x \in V$. □

Definition 12.224. *A topological space $\langle X, \mathcal{T}\rangle$ is normal if for every pair of disjoint closed sets A, B there exists a pair of disjoint open sets U, V with $A \subseteq U \wedge B \subseteq V$*

Theorem 12.225. *A normal topological space where every singleton is closed is regular, a regular space where every singleton is closed is Hausdorff*

Proof.

1. Let $\langle X, \mathcal{T}\rangle$ be a normal space so that $\forall x \in X$ we have $\{x\}$ is closed. If A is a closed set and $x \notin A$ then as $\{x\}$ is closed there exists by normality open sets U, V with $A \subseteq U, \{x\} \subseteq V \Rightarrow x \in V$ such that $U \bigcap V = \emptyset$ proving that $\langle X, \mathcal{T}\rangle$ is regular.
2. Let $\langle X, \mathcal{T}\rangle$ be a regular space so that $\forall x \in X$ we have $\{x\}$ is closed. If now $x, y \in X$ with $x \neq y \Rightarrow y \notin \{x\}$ a closed set so that by regularity there exists disjoint open sets U, V with $y \in U$ and $\{x\} \subseteq V \Rightarrow x \in V$ proving Hausdorff. □

Definition 12.226. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and $x \in X$ then a set A is a neighborhood of x iff there exists a open set U such that $x \in U \subseteq A$. If A is also a open set then it is called a open neighborhood.*

Definition 12.227. *Let $\langle X, \mathcal{T}\rangle$ be a topological space, $x \in X$ then a fundamental system of neighborhoods of x is a set $\mathcal{N}$ of neighborhoods of x such that for every neighborhood A of x there exists a $N \in \mathcal{N}$ such that $x \in N \subseteq A$*

Definition 12.228. *A topological space $\langle X, \mathcal{T} \rangle$ is first countable if every element of X has a countable fundamental system of neighborhoods.*

Theorem 12.229. *A metric space $\langle X, d \rangle$ is first countable*

Proof. Given $x \in X$ define the countable set $\mathcal{N}x = \left\{B_d\left(x, \frac{1}{n}\right) | n \in \mathbb{N}\right\}$ then if A is a neighborhood of x there exists a open set U with $x \in U \subseteq A$. So there exists a $\varepsilon \in \mathbb{R}_+$ such that $x \in B_d(x, \varepsilon)$. Using 9.55 there exists a $n \in \mathbb{N}$ such that $0 < \frac{1}{n} < \varepsilon$ so that $x \in B_d\left(x, \frac{1}{n}\right) \subseteq B_d(x, \varepsilon) \subseteq U \subseteq A$ so that $\mathcal{N}_x$ is a fundamental system of neighborhoods of x. $\square$

Definition 12.230. *A topological space $\langle X, \mathcal{T} \rangle$ is second countable if it has a countable basis*

Theorem 12.231. *A second countable topological space is first countable*

Proof. Let $\langle X, \mathcal{T} \rangle$ be a second countable topological space and let $\mathcal{B} \subseteq \mathcal{T}$ be the countable basis of $\mathcal{T}$ then if $x \in X$ and A a neighborhood of x then there exists a $U \in \mathcal{T}$ such that $x \in U \subseteq A$. As $\mathcal{B}$ is a basis there exists a $B \in \mathcal{B}$ such that $x \in B \subseteq U \subseteq A$. So $\mathcal{B}$ is a fundamentally system of neighborhoods of every element of x. $\square$

12.9 Toppological Vector Spaces

Definition 12.232. *A topological vector space is a vector space $\langle X, +, \cdot \rangle$ over $\mathbb{K}$ together with a topology $\mathcal{T}$ on X such that*

1. *$+: X \times X \rightarrow X$ defined by $(x, y) \rightarrow +(x, y) = x + y$ is continuous*
2. *$\cdot: F \times X \rightarrow X$ defined by $(\alpha, x) \rightarrow \cdot(\alpha, x) = \alpha \cdot x$ is continuous*

using the product topology of $X \times X$ and $\mathbb{K} \times X$ (using the toplogy of $\langle \mathbb{K}, \| \rangle$ and of X).

Example 12.233. Using 12.157 we have that a normed vector space is a topological vector space.

Definition 12.234. *Let X, Y be topological vector spaces then a function $L: X \rightarrow Y$ is a toplinear isomorphism iff*

1. *L is a linear isomorphism, meaning that*
   a. *L is a bijection*
   b. *$L(x + y) = L(x) + L(y)$*
   c. *$L(\alpha \cdot x) = \alpha \cdot L(x)$*
2. *L and L^{-1} are continuous (using the product topology on*

Theorem 12.235. *Let $\langle X, +, \cdot \rangle$ be a topological vector space then the functions*

1. *$+_x: X \rightarrow X$ defined by $y \rightarrow +_x(y) = x + y$*

2. $\cdot_\alpha: X \rightarrow X$ *defined by* $y \rightarrow \cdot_\alpha(y) = \alpha \cdot y$

are continuous.

Proof.

1. Let U be open and take $y \in (+_x)^{-1}(U)\{y \in X | x + y \in U\}$ then $(x, y) \in +^{-1}(U)$ and by continuity of $+$ (and the definition of the product topology) there exists open V_1, V_2 with $(x, y) \in V_1 \times V_2$ such that $+(V_1 \times V_2) \subseteq U \Rightarrow V_1 + V_2 \subseteq U$ and thus as $x \in V_1$ we have $+_x(V_2) = \{x\} + V_2 \subseteq V_1 + V_2 \subseteq U$ proving continuity of $+_x$.
2. Let U be open and take $x \in \cdot_\alpha(U) = \{x \in X | \alpha \cdot x \in U\}$ then $(\alpha, x) \in \cdot^{-1}(U)$ and by continuity of $\cdot$ there exists open V_1, V_2 with $(\alpha, x) \in V_1 \times V_2$ such that $\cdot(V_1 \times V_2) = V_1 \cdot V_2 \subseteq U$ and thus $\cdot_\alpha(V_2) = \{\alpha\} \cdot V_2 \subseteq V_1 \cdot V_2 \subseteq U$ proving continuity. □

12.10 Compact Spaces

Definition 12.236. *A topological space $\langle X, \mathcal{T}\rangle$ is compact iff for every family of $\{U_i\}_{i \in I}$ of open sets such that $X = \bigcup_{i \in I} U_i$ [$\{U_i\}_{i \in I}$ cover's X] there exists a finite $J \subseteq I$ such that $X = \bigcup_{i \in J} U_i$*

Definition 12.237. *Let $\langle X, \mathcal{T}\rangle$ be a topological space then a subset $C \subseteq X$ is compact in X iff $\langle C, \mathcal{T}_C\rangle$ is a compact topological space (where $\mathcal{T}_C$ is the subspace topology)*

Theorem 12.238. *Let $\langle X, \mathcal{T}\rangle$ be a topological space then a subset $C \subseteq X$ is compact in X if and only if for every family $\{U_i\}_{i \in I}$ of open sets such that $C \subseteq \bigcup_{i \in I} U_i$ there exists a finite $J \subseteq I$ such that $C \subseteq \bigcup_{i \in J} U_i$*

Proof.

1. **($\Rightarrow$)** If $\{U_i\}_{i \in I}$ is a collection of open sets such that $C \subseteq \bigcup_{i \in I} U_i$ then $C = \bigcup_{i \in I} (C \bigcap U_i)$ where $\forall i \in I$ we have that $U_i \bigcap C \in \mathcal{T}_C$ so that by the fact that $\langle C, \mathcal{T}_C\rangle$ is compact there exists a finite $J \subseteq I$ such that $C = \bigcup_{i \in J} (U_i \bigcap C) \Rightarrow C \subseteq \bigcup_{i \in J} U_i$
2. **($\Leftarrow$)** Assume that $\{V_i\}_{i \in I}$ is a collection of open sets in $\mathcal{T}_C$ such that $C = \bigcup_{i \in I} V_i$ then by the definition of the subspace topology we have $\forall i \in I$ there exists a U_i such that $V_i = U_i \bigcap C$ so that $C = \bigcup_{i \in I} (U_i \bigcap C) \subseteq \bigcup_{i \in I} U_i$ so by the hypothesis there exists a finite $J \subseteq I$ such that $C \subseteq \bigcup_{i \in J} U_i \Rightarrow C = \bigcup_{i \in J} (U_i \bigcap C) = \bigcup_{i \in J} V_i$ proving that $\langle C, \mathcal{T}_C\rangle$ is a compact topological space. □

Example 12.239. If $\langle X, \mathcal{T}\rangle$is a topological space then $\emptyset$ is a compact subset of X

Proof. If $\{U_i\}_{i \in I}$ is a non empty family of open sets covering $\emptyset$ then we have for I either

$\boldsymbol{I = \emptyset}$**.** then $J = \emptyset = I$ is finite so $\{U_i\}_{i \in J}$ is a finite covering of $\emptyset$

$I \neq \emptyset$. then $\exists i \in I$ if we take $J = \{i\} \subseteq I$ then J is finite and $\emptyset \subseteq E_i = \bigcup_{j \in \{i\}} U_i$ □

Theorem 12.240. *Let $\langle X, \mathcal{T} \rangle$ be a topological space and $\{C_i\}_{i \in \{1, \ldots, n\}}$ be a finite family of compact subsets of X then $\bigcup_{i \in \{1, \ldots, n\}} C_i$ is compact.*

Proof. Let $\bigcup_{i \in I} U_i$ is a open cover of $\bigcup_{i \in \{1, \ldots, n\}} C_i$ then $\forall i \in \{1, \ldots, n\}$ we have $C_i \subseteq \bigcup_{j \in I} U_j \underset{C_i \text{is compact}}{\Rightarrow} \exists \text{finite } I_i \subseteq I \vdash C_i \subseteq \bigcup_{j \in I_i} U_j$ so that $\bigcup_{i \in \{1, \ldots, n\}} C_i \subseteq \bigcup_{i \in \{1, \ldots, n\}} (\bigcup_{j \in I_i} U_i)$. If now $x \in \bigcup_{i \in \{1, \ldots, .n\}} (\bigcup_{j \in I_i} U_i)$ there exists a $i \in \{1, \ldots, n\}$ such that $x \in \bigcup_{j \in I_i} U_i \Rightarrow \exists i \in \{1, \ldots, n\}$ such that $\exists j \in I_i$ with $x \in U_j \Rightarrow \exists j \in \bigcup_{i \in \{1, \ldots, n\}} I_i \subseteq I \vdash x \in U_j$ proving that $\bigcup_{i \in \{1, \ldots, n\}} (\bigcup_{j \in I_i} U_j) \subseteq \bigcup_{j \in \bigcup_{i \in \{1, \ldots, n\}} I_i} U_j$ and thus we have $\bigcup_{i \in \{1, \ldots, n\}} C_i \subseteq \bigcup_{j \in \bigcup_{i \in \{1, \ldots, n\}} I_i} U_i$ where $\bigcup_{i \in \{1, \ldots, n\}} I_i$ is a finite subset of I. □

Theorem 12.241. *Let $\langle X, \mathcal{T} \rangle$ be a topological space and let $A \subseteq X$ then if $C \subseteq A$ is compact in $\langle A, \mathcal{T}_A \rangle$ iff it is compact in $\langle X, \mathcal{T} \rangle$*

Proof.

- **$\Rightarrow$.** Let C be compact in $\langle X, \mathcal{T} \rangle$. Let $\{U_i\}_{i \in I}$ be a set of open sets in $\mathcal{T}$ with $C \subseteq \bigcup_{i \in I} U_i \Rightarrow C = C \bigcap A \subseteq (\bigcup_{i \in I} U_i) \bigcap A = \bigcup_{i \in I} (U_i \bigcap A) \underset{C \text{is compact in} \langle A, \mathcal{T}_A \rangle}{\Rightarrow} \exists J \subseteq I$ with J finite and $C \subseteq \bigcup_{i \in J} (U_i \bigcap A) \subseteq \bigcup_{i \in J} U_i$ so that C is compact in $\langle X, \mathcal{T} \rangle$
- **$\Leftarrow$.** Let C be compact in $\langle A, \mathcal{T}_A \rangle$ then if $C \subseteq \bigcup_{i \in I} U_i$ (a open cover of A in $\mathcal{T}$) then $C = C \bigcap A \subseteq \bigcup_{i \in I} (U_i \bigcap A)$ (a open cover of A in $\mathcal{T}_A$), so by compactness we have that there exists a finite set $J \subseteq I$ such that $C \subseteq \bigcup_{i \in J} (U_i \bigcap A) \subseteq \bigcup_{i \in J} U_i$ proving that C is compact in $\mathcal{T}$. □

Theorem 12.242. *Let $\langle X, d \rangle$ be a metric space and let $C \subseteq X$ be a compact set then C is bounded (see 12.65)*

Proof. Let C be a compact subset of X then $\forall c \in C$ take $B_d(c, 1) \ni c$ which is open. So we have $C \subseteq \bigcup_{c \in C} B_d(c, 1)$ and thus by compactness there exists a finite set $\{c_1, \ldots, c_n\} \subseteq C$ such that $C \subseteq \bigcup_{i \in \{1, \ldots, n\}} B_d(c_i, 1)$. Take then $M = 2 + \max(\{d(c_i, c_j) | (i, j) \in \{1, \ldots, n\} \times \{1, \ldots, n\}\})$ (which exists because of 5.50 and 5.44). If now $x, y \in C \Rightarrow \exists i, j \in \{1, \ldots, n\}$ such that $x \in B_d(c_i, 1)$ and $y \in B_d(c_j, 1)$ and thus $d(x, y) \leqslant d(x, c_i) + d(c_i, y) \leqslant d(x, c_i) + d(c_i, c_j) + d(c_j, y) < 1 + d(c_i, c_j) + 1 = 2 + d(c_i, c_j) \leqslant M$ proving that C is bounded. □

Theorem 12.243. *Let $\langle X, \mathcal{T}_X \rangle$, $\langle Y, \mathcal{T}_Y \rangle$ be topological spaces and $f: X \rightarrow Y$ a continuous map then if $C \subseteq X$ is compact then $f(C)$ is compact.*

Proof. Let $\{V_i\}_{i \in I}$ be a family of open sets in Y so that $f(C) \subseteq \bigcup_{i \in I} V_i$ then $\forall x \in C$ we have $f(x) \in f(C)$ and thus $\exists i \in I$ such that $f(x) \in V_i$ so that by continuity of f there exists a U_i open in X with $x \in U_i$ and $f(x) \in f(U_i) \subseteq V_i$ so that $C \subseteq \bigcup_{i \in I} U_i$. By compactness of C there exist a finite $J \subseteq I$ so that $C \subseteq \bigcup_{i \in J} U_i \underset{2.54}{\Rightarrow} f(C) \subseteq f(\bigcup_{i \in J} U_i) \underset{2.58}{=} \bigcup_{i \in J} f(U_i) \subseteq \bigcup_{i \in J} V_i$ proving that $f(C)$ is compact. □

Theorem 12.244. *Let $\langle X, \mathcal{T}\rangle$ be a Hausdorff topological space then every compact subspace is closed*

Proof. Let C be a compact subset of X then if $y \in X \setminus C$ we have that $\forall x \in C$ there exists open sets U_x, V_x such that $x \in U_x$, $y \in V_x$ and $U_x \bigcap V_x \neq \emptyset$. So $C \subseteq \bigcup_{x \in X} U_x$ and by compactness there exists a finite set $F \subseteq X$ such that $C \subseteq \bigcup_{x \in F} U_x$. Now $y \in \bigcap_{z \in F} V_z$ which is open (because F is finite) and $(\bigcap_{z \in F} V_z) \bigcap C \subseteq (\bigcap_{z \in F} V_z) \bigcap (\bigcup_{x \in F} U_x) = \bigcup_{x \in F} ((\bigcap_{z \in F} V_z) \bigcap U_x) \subseteq \bigcup_{x \in F} (V_x \bigcap U_x) = \bigcup_{x \in F} \emptyset = \emptyset$ so that $y \in \bigcap_{z \in F} V_x \subseteq X \setminus C$ proving that $X \setminus C$ is open and thus that C is closed. □

Theorem 12.245. *Let $\langle X, \mathcal{T}\rangle$ be a topological space then every closed subset of a compact subset of X is compact.*

Proof. Let $C \subseteq X$ be compact and let $A \subseteq C$ be a closed set and let $\{U_i\}_{i \in I}$ be a family of open sets covering A (so $A \subseteq \bigcup_{i \in I} U_i$) then as $C = A \bigcup (C \setminus A) \subseteq A \bigcup (X \setminus A)$ we have that if $\mathcal{W} = \{U_{i \in I} | i \in I\} \bigcup \{X \setminus A\}$ then $\{W\}_{W \in \mathcal{W}}$ covers C. By compactness we have that a finite subset $\mathcal{V}$ of $\mathcal{W}$ covers C and thus A so $A \subseteq \bigcup_{U \in \mathcal{V}} U$. If now $A \nsubseteq \bigcup_{U \in \mathcal{V} \setminus \{X \setminus A\}} U$ then there exists a $a \in A$ with $a \in \bigcup_{U \in \mathcal{V} \setminus \{X \setminus A\}} U$ which as $A \subseteq \bigcup_{U \in \mathcal{V}} U$ means that $a \in X \setminus A$ contradicting that $a \in A$ so we must have that $A \subseteq \bigcup_{U \in \mathcal{V} \setminus \{X \setminus A\}} U$. If we define now $\{\mathcal{I}_U\}_{U \in \mathcal{V} \setminus \{X \setminus A\}}$ by $\mathcal{I}_U = \{i \in I | U_i = U\}$ then as $\mathcal{V} \setminus \{X \setminus A\} \subseteq \mathcal{W} \setminus \{X \setminus A\} = \{U_i | i \in I\}$ we have that $\mathcal{I}_U \neq \emptyset$. Using the axiom of choice (see 2.201) we have then that there exist a function c: $\mathcal{V} \setminus \{X \setminus A\} \rightarrow \bigcup_{U \in \mathcal{V} \setminus \{X \setminus A\}} \mathcal{I}_U \subseteq I$ with $c(U) \in \mathcal{I}_U$. So we have then the surjection $c: \mathcal{V} \setminus \{X \setminus A\} \rightarrow c(\mathcal{V} \setminus \{X \setminus A\}) \subseteq I$ which means that because $\mathcal{V} \setminus \{X \setminus A\}$ is finite that $c(\mathcal{V} \setminus \{X \setminus A\})$ is finite. Now if $a \in A$ then as $A \subseteq \bigcup_{U \in \mathcal{V} \setminus \{X \setminus A\}} U$ there exist a $U \in \mathcal{V} \setminus \{X \setminus A\}$ such that $a \in U \underset{c(U) \in \mathcal{I}_U}{=} U_{c(U)} \subseteq \bigcup_{i \in c(\mathcal{V} \setminus \{X \setminus A\})} U_i$ proving that $A \subseteq \bigcup_{i \in c(\mathcal{V} \setminus \{X \setminus A\})} U_i$ and thus that A is compact. □

Lemma 12.246. *Given $\langle \mathbb{R}, \|\rangle$ the normed space with the canonical norm $||$, S a none empty closedset such that* $\sup(S)$ *[or* $\inf(S)$*] exists then* $\sup(S) \in S$ *[*$\inf(S) \in S$*]*

Proof. First if $\sup(S) \in U \in \mathcal{T}_{||}$ then there exists a $\varepsilon > 0$ such that $\sup(S) \in B_{||}(\sup(S), \varepsilon) \subseteq U$. As $\sup(S) - \varepsilon < \sup(S)$ there exists using the definition of the supremum a $y \in S$ such that $\sup(S) - \varepsilon < y \leqslant \sup(S) \Rightarrow y \in B_{||}(\sup(S), \varepsilon) \subseteq U$ proving that $U \bigcap S = \emptyset$ hence by 12.21 and the closure of S we have $\sup(S) \in S$.

Second if $\inf(S) \in U \in \mathcal{T}_{||}$ then there exists a $\varepsilon > 0$ such that $\inf(S) \in B_{||}(\inf(S), \varepsilon) \subseteq U$. As $\inf(S) - \varepsilon < \inf(S)$ there exists using the definition of the infinum a $y \in S$ such that $\inf(S) \leqslant \inf(S) + \varepsilon \Rightarrow y \in B_{||}(\inf(S), \varepsilon) \subseteq U$ proving that $U \bigcap S = \emptyset$ hence by 12.21 and the closure of S we have $\inf(S) \in S$. □

Theorem 12.247. (Extreme value Theorem) *Let $\langle X, \mathcal{T}\rangle$ be a topologicval space, C a non empty compact subset and $f: X \rightarrow \mathbb{R}$ a continuous function then there exists a $m, M \in X$ such that $\forall x \in X$ we have $f(m) \leqslant f(x) \leqslant f(M)$*

Proof. As X is compact we have by 12.243 that $f(C)$ is a compact set in $\mathbb{R}$, so by 12.242 $f(C)$ is bounded. Hence there exissts a M such that $\forall x, y \in f(C)$ we have $|f(x) - f(y)| \leqslant M$. As $C \neq \emptyset$ there exists a $x_0 \in C$ so that $\forall x \in C$ we have $|f(x) - f(x_0)| \leqslant M \Rightarrow f(x_0) - M \leqslant f(x) \leqslant f(x_0) + M$. Using the conditional completeness of the reals we have that $\inf(f(C))$ and $\sup(f(C))$ exists. Using the previous theorem we have then that $\inf(f(C)), \sup(f(C)) \in f(C)$ so that $\exists m$, $M \in C$ such that $f(m) = \inf(f(C)) \wedge f(M) = \sup(f(C))$. Hence if $x \in C$ then $f(x) < f(C)$ and thus $f(m) = \inf(f(C)) \leqslant f(x) \leqslant \sup(f(C)) = f(M)$. □

Theorem 12.248. *Let $\langle X, \mathcal{T}\rangle$ be a Hausdorff compact topological space then it is regular and normal.*

Proof.

1. **(regular)** Let A be a closed set and $x \in X$ with $x \notin A$ then because of Hausdorff we have that $\forall a \in A$ we have the existence of disjoint open sets V_a, U_a with $x \in V_a, a \in U_a$ and $V_a \bigcap U_a = \emptyset$ then $A \subseteq \bigcup_{a \in A} U_a$ and because A is compact (closed sets in a compact space are compact) there exist a finite $B \subseteq A$ such that $A \subseteq \bigcup_{a \in B} U_a$. We have then that $V = \bigcap_{a \in B} V_a$ is a open set (because B is finite), $a \in V$ and $V \bigcap A = (\bigcap_{b \in B} V_b) \bigcap A \subseteq (\bigcap_{b \in B} V_b) \bigcap (\bigcup_{a \in B} U_a) = \bigcup_{a \in B} ((\bigcap_{b \in B} V_b) \bigcap U_a) \subseteq \bigcup_{a \in B} (V_a \bigcap U_a) = \emptyset \Rightarrow V \bigcap A = \emptyset$ proving regularity.
2. **(normal)** Let A, B be two disjoint closed sets then $\forall a \in A$ we have using (1) the existence of disjoint open sets V_a, U_a so that $B \subseteq V_a$ and $a \in U_a$ so that $A \subseteq \bigcup_{a \in A} U_a$. From the compactness of A (because closed sets in a compact space are compact) there exist a finite set $F \subseteq A$ so that $A \subseteq \bigcup_{a \in F} U_a$ then $B \subseteq \bigcap_{a \in F} V_a$ and if we take $U = \bigcup_{a \in F} U_a$ (which is open) and $V = \bigcap_{a \in F} V_a$ (which is open as F is finite) then $A \subseteq U$, $B \subseteq V$ and $V \bigcap U = (\bigcap_{b \in B} V_b) \bigcap (\bigcup_{a \in B} U_a) = \bigcup_{a \in B} ((\bigcap_{b \in B} V_b) \bigcap U_a) \subseteq \bigcup_{a \in B} (V_a \bigcap U_a) = \emptyset \Rightarrow U \bigcap V = \emptyset$ proving normality. □

Theorem 12.249. *Let $\langle X, \mathcal{T}\rangle$ be a topological space with a basis $\mathcal{B}$ then a subset $C \subseteq X$ is compact iff every cover of C by sets in the base contains a finite cover.*

Proof.

1. **($\Rightarrow$)** This is trivial from $\mathcal{B} \subseteq \mathcal{T}$
2. **($\Leftarrow$)** Let $\{U_i\}_{i \in I}$ be a family of open sets such that $C \subseteq \bigcup_{i \in I} U_i$ then $\forall x \in C$ there exists a $i_x \in I$ such that $x \in U_{i_x}$ and thus there exist a $B_x \in \mathcal{B}$ such that $x \in B_x \subseteq U_x$. So we have $C \subseteq \bigcup_{x \in C} B_x$ and by the hypothesis there exists a finite $\{x_1, \ldots, x_n\} \subseteq C$ such that $C \subseteq \bigcup_{x \in \{x_1, \ldots, x_n\}} B_x \underset{B_x \subseteq U_{i_x}}{\Rightarrow} \bigcup_{x \in \{x_1, \ldots, x_n\}} U_{i_x} = \bigcup_{i \in \{i_{x_1}, \ldots, i_{x_n}\}} U_{i_x}$ proving as $\{i_{x_1}, \ldots, i_{x_n}\}$ is finite that C is compact. □

Definition 12.250. *Let $\langle X, \mathcal{T}\rangle$ be a topological space then a subset C is called sequential compact if every infinite $B \subseteq C$ has a limit point (see 12.18) that is in C.*

Theorem 12.251. *A compact subspace of a topological space $\langle X, T\rangle$ is sequential compact*

Proof. We proceed by contradiction. Let C be a compact subset and let $B \subseteq C$ be a infinite subset that does not have a limit point in C. Then $\forall x \in C$ there exists a open U_x with $x \in U_x$ and $U_x \bigcap (B \setminus \{x\}) = \emptyset$. If $y \in U_x \bigcap B$ then if $y \neq x \Rightarrow y \notin \{x\} \Rightarrow y \in (U_x \bigcap B) \setminus \{x\} = U_x \bigcap (B \setminus \{x\})$ contradicting $U_x \bigcap (B \setminus \{x\}) = \emptyset$ so we must have $y = x$ or $U_x \bigcap B \subseteq \{x\}$. Now $C \subseteq \bigcup_{x \in C} U_x$ and as C is compact there exists a $F \subseteq C$ such that $C \subseteq \bigcup_{x \in C} U_x$ and thus $B \underset{B \subseteq C}{=} B \bigcap C = B \bigcap (\bigcup_{x \in F} U_x) = \bigcup_{x \in F} (B \bigcap U_x) \subseteq \bigcup_{x \in F} \{x\} = F$ which is finite contradicting that B is infinite. □

Theorem 12.252. *Let $\langle X, d\rangle$ be a metric space equipped with the metric topology. Let $K \subseteq X$ be a sequential compact subset of X and $\{U_i\}_{i \in I}$ be a open cover of K then $\exists \delta \in \mathbb{R}_+$ such that $\forall x \in K$ there $\exists i_x \in I$ such that $x \in B_d(x, \delta) \subseteq U_{i_x}$*

Proof. We prove this by contradiction. So assume that $\forall \delta \in \mathbb{R}_+$ there $\exists x \in K$ such that $\forall i \in I$ we have $B_d(x, \delta) \nsubseteq U_i$. In particular this means that (using the isomorphism $i_{\mathbb{Q}_{\mathbb{N}_0}}: \mathbb{N}_0 \rightarrow \mathbb{N}_{0\,\mathbb{R}} \subseteq \mathbb{R}$ (see 8.56) and noting $i_{Q_{\mathbb{N}_0}}(n)$ by n) we have $\forall n \in \mathbb{N}$ there exists a $x_n \in K$ such that $\forall i \in I$ we have $B_d\left(x_n, \frac{1}{n}\right) \nsubseteq U_i$. This defines a sequence $\{x_i\}_{i \in \mathbb{N}}$ of elements in K such that $B_d\left(x_n, \frac{1}{n}\right) \nsubseteq U_i$ $\forall i \in I, \forall n \in \mathbb{N}$. Define now the set $O = \{x_n | n \in \mathbb{N}\} \subseteq K$. If this set was finite then using 5.52 there exists a $x_N \in O$ such that $\forall n \in \mathbb{N}$ there exists a $m \in \mathbb{N}$ with $m \geqslant n$ and $x_m = x_N$. Now as $x_N \in O \subseteq K \underset{\{U_i\}_{i \in I} \text{is a open cover of } K}{\Rightarrow} \exists i \in I$ with $x_N \in U_i \underset{U_i \text{is open}}{\Rightarrow} \exists \varepsilon \in \mathbb{R}_+$ such that $x_N \subseteq B_d(x_N, \varepsilon) \subseteq U_i$. Take then $M \in \mathbb{N}_{0\,\mathbb{R}}$ such that $0 < \frac{1}{M} < \varepsilon$ (see 9.55) then by the above there exists a $m \in \mathbb{N}$ with $m \geqslant M$ (and thus $\frac{1}{M} \geqslant \frac{1}{m}$) so that $x_N = x_m$, from this it follows that $x_m \in B_d\left(x_m, \frac{1}{m}\right) = B_d\left(x_N, \frac{1}{m}\right) \subseteq B_d\left(x_N, \frac{1}{M}\right) \subseteq B_d(x_n, \varepsilon) \subseteq U_i$ contradicting the fact that $B_d\left(x_m, \frac{1}{m}\right) \nsubseteq U_i$ $\forall i \in I$. We must thus conclude that O is infinite. As K is sequential compact we have thus a limit point $k \in K$ for O. Given that $K \subseteq \bigcup_{i \in I} U_i$ there exists a $i_0 \in I$ such that $k \in U_{i_0} \underset{U_{i_0} \text{is open}}{\Rightarrow} \exists \varepsilon_0 \in \mathbb{R}_+$ such that $k \in B_d(k, \varepsilon_0) \subseteq U_{i_0}$. The set $P = \left\{n \in \mathbb{N} | x_n \in B_d\left(k, \frac{\varepsilon_0}{2}\right)\right\}$ is a infinite set of natural numbers [if it was finite we could take $\varepsilon = \min\left(\frac{\varepsilon_0}{2}, \min(\{d(k, x_i) | i \in P\})\right)$ [$\min(\{d(k, x_i) | i \in P\})$ exists if P is finite] then $k \in B_d(k, \varepsilon) \subseteq B_d\left(k, \frac{\varepsilon_0}{2}\right)$ and $\forall i \in \mathbb{N}$ we have either $i \notin P$ so that then $x_i \notin B_d\left(k, \frac{\varepsilon_0}{2}\right) \supseteq B_d(k, \varepsilon) \Rightarrow x_i \notin B_d(k, \varepsilon)$ or $i \in P$ and then as $d(k, x_i) \geqslant \varepsilon$ we have $x_i \notin B_d(k, \varepsilon)$. So $\forall i \in \mathbb{N}$ we have $x_i \notin B_d(k, \varepsilon) \Rightarrow O \bigcap B_d(k, \varepsilon) = \emptyset$ contradicting the fact that k is the limit point of O. So we must have that P is infinite]. As $P \subseteq \mathbb{N}$ is infinite there exists a $m \in P$ such that $\frac{1}{m} < \frac{\varepsilon_0}{2}$ [using 9.55 there exist a $N \in \mathbb{N}$ such that $\frac{1}{N} < \frac{\varepsilon_0}{2}$ and as P is infinite we have that $P \nsubseteq S_N = \{1, \ldots, N-1\}$ so that $\exists m \in P$ with $m \in \{1, \ldots, N-1\} \Rightarrow m > N-1 \Rightarrow m \geqslant N \Rightarrow \frac{1}{m} \leqslant \frac{1}{N} < \frac{\varepsilon_0}{2}$], as $m \in P$ we have that $x_m \in B_d\left(k, \frac{\varepsilon_0}{2}\right) \Rightarrow d(k, x_m) < \frac{\varepsilon_0}{2}$. Now if $z \in B_d\left(x_m, \frac{1}{m}\right) \Rightarrow d(x_m, z) < \frac{1}{m} \Rightarrow d(k, z) \leqslant d(k, x_m) + d(x_m, z) < \frac{\varepsilon_0}{2} + \frac{1}{m} < \frac{\varepsilon_0}{2} + \frac{\varepsilon_0}{2} = \varepsilon_0 \Rightarrow z \in B_d(k, \varepsilon_0) \Rightarrow B_d\left(x_m, \frac{1}{m}\right) \subseteq B_d(k, \varepsilon_0) \subseteq U_{i_0}$ contradicting the fact that $\forall i \in I$ we have $\forall n \in \mathbb{N}$ that $B_d\left(x_n, \frac{1}{n}\right) \nsubseteq U_i$. □

Theorem 12.253. *Let $\langle X, d\rangle$ be a metric space with the metric topology and $K \subseteq X$ then the following are equivalent*

1. *K is compact*

2. *K is sequential compact*

Proof.

1. (⇒) This is already proved in 12.251

2. (⇐) Let K be sequential compact and $\{U_i\}_{i\in I}$ a cover of K. If $K=\emptyset$ then it is trivially compact so lets assume that $K\neq\emptyset$. Using 12.252 there exists a $\delta\in\mathbb{R}_+$ such that $\forall x\in K$ there exists a $i\in I$ such that $x\in B_d(x,\delta)\subseteq U_i$. If we prove now that

$$\exists F\,(\text{finite})\subseteq K\,\text{such that}\,K\subseteq\bigcup_{x\in F}B_d(x,\delta) \tag{12.23}$$

then $\forall f\in F$ there exists a $j_f\in I$ such that $x_{j_f}\in B_d(x_{j_f},\delta)\subseteq U_{j_f}$ defining a function $\sigma\colon F\rightarrow I$ given by $f\rightarrow j_f$ (by the Axiom of Choice 2.202) so that $K\subseteq\bigcup_{f\in F}B_d(x_{\sigma(f)},\delta)\subseteq\bigcup_{f\in F}U_{\sigma(f)}=\bigcup_{j\in P}U_j$ where $P=\sigma(F)\subseteq I$ is finite, proving compactness. We prove 12.23 by contradiction. So assume that

$$\forall F\,(\text{finite})\subseteq K\,\text{we have that}\,K\nsubseteq\bigcup_{x\in F}B_d(x,\delta) \tag{12.24}$$

Given a family $\{x_i\}_{i\in\{0,\ldots,n\}}$ we have that the set $\{x_i|i\in\{0,\ldots,n\}\}$ is finite so that there exists a $\rho(\{x_i\}_{i\in\{0,\ldots,n\}})\in K\setminus\left(\bigcup_{i\in\{0,\ldots,n\}}B_d(x_i,\delta)\right)$, using the Axiom of Choice (see 2.202) this defines a function $\rho\colon\mathcal{M}=\{\{x_i\}_{i\in\{0,\ldots,n\}}|n\in\mathbb{N}_0\wedge\forall i\in\{0,\ldots n\}\text{ we have }x_i\in K\}\rightarrow K$. As $K\neq\emptyset$ there exists a $x_0\in K$ so that using recursion (see 5.21) there exists a sequence $f\colon\mathbb{N}_0\rightarrow K$ such that

a. $f(0)=x_0$

b. $f(n+1)=\rho(\{f(i)_i\}_{i\in\{0,\ldots,n\}})\in K\setminus\left(\bigcup_{i\in\{0,\ldots,n\}}B_d(f(i),\delta)\right)$

Or if $n\in\mathbb{N}$ then $x_n\notin K\setminus\left(\bigcup_{i\in\{0,\ldots,n-1\}}B_d(f(i),\delta)\right)$. If now $i\neq j$ then we have $i<j$ (for the other case just switch i and j) then $f(i)\notin B_d(f(j),\delta)\Rightarrow d(f(i),f(j))\geqslant\delta>0\Rightarrow f(i)\neq f(j)$ proving that f is injective. Using 5.49 we have then that $O=f(\mathbb{N}_0)\subseteq K$ is finite so it has a accumulation point $x\in K$ (because of sequential compactness). Take then $B_d\left(x,\frac{\delta}{2}\right)$ then there exists a $i\in\mathbb{N}_0$ such that $f(i)\neq x\Rightarrow d(x,f(i))>0$ and $f(i)\in B_d\left(x,\frac{\delta}{2}\right)$, if we take then $\varepsilon=\min\left(\frac{\delta}{4},\frac{d(x,f(i))}{2}\right)$ then $d(x,f(i))>\varepsilon\Rightarrow f(i)\notin B_d(x,\varepsilon)$. If $m\neq i$ and $f(m)\in B_d(x,\varepsilon)\underset{f(i)\notin B_d(f(m),\delta)\text{ or }f(m)\notin B_d(f(i),\delta)}{=}\delta\leqslant d(f(m),f(i))\leqslant d(f(m),x)+d(x,f(i))<\varepsilon+\frac{\delta}{2}<\frac{\delta}{2}+\frac{\delta}{2}=\delta$ giving the contradiction $\delta<\delta$ so we must have that $f(m)\in B_d(x,\varepsilon)$ this together with $f(i)\in B_d(x,\varepsilon)$ proves that $O=f(\mathbb{N}_0)\bigcap B_d(x,\varepsilon)=\emptyset$ contradicting the fact that x is a accumulation point of O. So 12.24 is false proving 12.23 and thus the theorem. □

Definition 12.254. *Let X be a set then $K\subseteq\mathcal{P}(X)$, a set of subsets of X, is a **compact class** if for any sequence $\{K_i\}_{i\in\mathbb{N}}$ of elements in K with $\bigcap_{i\in\mathbb{N}}K_i=\emptyset$ we have that $\exists N\in\mathbb{N}$ such that $\bigcap_{i\in\{1,\ldots,N\}}K_i=\emptyset$*

Example 12.255. Let $\langle X, \mathcal{T}\rangle$ be a Hausdorff topological space and K a set of compact sets in X then K is a countable compact class.

Proof. We prove this by contradiction so let $\{K_i\}_{i\in\mathbb{N}}$ a countable family of elements in K with $\bigcap_{i\in\mathbb{N}} K_i = \emptyset$ such that $\forall n \in \mathbb{N}$ we have $\bigcap_{i\in\{1,\ldots,n\}} K_i \neq \emptyset$. Define then $\{E_i\}_{i\in\mathbb{N}}$ by $E_i = \bigcap_{j=1}^{i} K_j$ then as K_j is closed (see 12.244) we have that $\{X \setminus E_i\}_{i\in\mathbb{N}}$ is a family of open sets. If now $x \in K_1$ then as $\bigcap_{i\in\mathbb{N}} K_i = \emptyset$ there exists a $n\in\mathbb{N}$ such that $x \notin K_n \Rightarrow x \in X \setminus K_n$ which proves that $K_1 \subseteq \bigcup_{i\in\mathbb{N}} (K \setminus E_i)$, using compactness of K_1 there exists then a finite $\{i_1, \ldots, i_n\} \subseteq \mathbb{N}$ such that $K_1 \subseteq \bigcup_{i\in\{i_1,\ldots,i_n\}} (K \setminus E_i)$. Take now $m = \max(i_{1,\ldots,} i_n)$ then as $\forall j \in \{i_1, \ldots, i_n\}$ we have $E_m \subseteq E_j \Rightarrow K \setminus E_j \subseteq K \setminus E_m$ so that $K_1 \subseteq K \setminus E_m$. As we assume that $E_m = \bigcap_{j=1}^{m} K_i \neq \emptyset$ there exists a $y \in E_m \subseteq K_1$ so that $y \in K_1 \subseteq K \setminus E_m \Rightarrow y \in E_m \wedge y \notin E_m$ a contradiction. So we must have that there exists a $N \in \mathbb{N}$ with $\bigcap_{i\in\{1,\ldots,N\}} K_i = \emptyset$. □

Theorem 12.256. *Let $[a_i, b_i]$, $a_i, b_i \in \mathbb{R}, i \in \mathbb{N}_0$ $a_i \leqslant b_i$ be a decreasing sequence [so $[a_{i+1}, b_{i+1}] \subseteq [a_i, b_i]$] of closed bounded intervals in $\mathbb{R}$ then*

1. *$\exists a, b \in \mathbb{R}, a \leqslant b$ such that $\bigcap_{i\in\mathbb{N}_0} [a_i, b_i] = [a, b]$*
2. *If the sequence $\{b_i - a_i\}_{i\in\mathbb{N}_0}$ has the limit 0 (see 12.312 for a definition of a limit) then there exists a number a such that $\bigcap_{i\in\mathbb{N}_0} [a_i, b_i] = \{a\}$*

Proof. First we prove by induction that $\forall i, j \in \mathbb{N}_0$ with $i \leqslant j$ we have $a_i \leqslant a_j \leqslant b_j \leqslant b_i$ so let $S_i = \{n \in \mathbb{N}_0 | a_i \leqslant a_{i+n} \leqslant b_{i+n} \leqslant b_i\}$ then we have

1. if $n = 0$ then $a_i \leqslant a_{i+0} \leqslant b_{i+0} \leqslant b_i$ proving that $0 \in S$
2. If $n \in \mathbb{N}_0$ then $[a_{i+(n+1)}, b_{i+(n+1)}] \subseteq [a_{i+n}, b_{i+1}] \underset{n\in S}{\Rightarrow} a_i \leqslant a_{i+n} \leqslant a_{i+(n+1)} \leqslant b_{i+(n+1)} \leqslant b_{i+n} \leqslant b_i$ proving that $n+1 \in S$

Using induction we have then that $S = \mathbb{N}_0$ so if $i \leqslant j$ then $j - i \in \mathbb{N}_0 = S \Rightarrow a_i \leqslant a_{i+(j-i)} = a_j \leqslant b_j = b_{i+(j-i)} \leqslant b_i$. Specially we have that $\forall i \in \mathbb{N}_0$ that $a_0 \leqslant a_i \leqslant b_i \leqslant b_0$.

1. $A = \{a_i | i \in \mathbb{N}_0\}$ is then bounded above by b_0 and $B = \{b_i | i \in \mathbb{N}\}$ is bounded below by a_0 so using 9.43 and 2.176 there exists a $a = \sup(A)$ and $b = \inf(B)$. As $\forall i, j \in \mathbb{N}_0$ we have either $i \leqslant j$ so that $a_i \leqslant a_j \leqslant b_j$ or $j \leqslant i$ then $a_i \leqslant b_i \leqslant b_j$ so in all case we have $a_i \leqslant b_j$ and thus $a_i \leqslant \inf(B)$ and $a = \sup(A) \leqslant \inf(B) = b \Rightarrow a \leqslant b$. And $\forall i \in \mathbb{N}_0$ we have by the definition of inf and sup that $a_i \leqslant a \leqslant b \leqslant b_i$ so that $[a, b] \subseteq \bigcap_{i\in\mathbb{N}_0} [a_i, b_i]$. Now assume that $x \in \bigcap_{i\in\mathbb{N}_0} [a_i, b_i]$ then $\forall i \in \mathbb{N}_0$ we have $a_i \leqslant x \leqslant b_i$ so x is a upper bound for A and a lower bound for B so that $a \leqslant x \leqslant b \Rightarrow x \in [a, b]$ proving that $\bigcap_{i\in\mathbb{N}_0} [a_i, b_i] = [a, b]$
2. By (1) we have that there exists a $a, b \in \mathbb{R}$ with $\bigcap_{i\in\mathbb{N}_0} [a_i, b_i] = [a, b]$. Suppose now that $a < b$ then $\forall i \in \mathbb{N}_0$ we have $a_i \leqslant a < b \leqslant b_i \Rightarrow 0 < b - a \leqslant b_i - a \leqslant b_i - a_i$. Now as by hypothesis we have that $\{b_i - a_i\}_{i\in\mathbb{N}_0}$ has limit 0 then given $\varepsilon = b - a > 0$ choose $N \in \mathbb{N}_0$ such that $\forall i \geqslant \mathbb{N}_0$ we have $b_i - a_i = |b_i - a_i| < \varepsilon \Rightarrow \varepsilon = b - a \leqslant b_N - a_N < \varepsilon$ giving the contradiction $\varepsilon < \varepsilon$ so we must have that $a = b$ but then $\bigcap_{i\in\mathbb{N}_0} [a_i, b_i] = [a, a] = \{a\}$ □

Theorem 12.257. *Let $[a, b], a \leqslant b, a, b \in \mathbb{R}$ be a closed and bounded interval in $\mathbb{R}$ then $[a, b]$ is a compact subset in $\mathbb{R}$ (using the norm topology generated by $||$)*

Proof. The proof is divided in two cases as $a \leqslant b$

1. $(\boldsymbol{a = b})$ then $[a, b] = \{a\}$ and if $\bigcup_{i \in I} U_i$ is a covering of $\{a\}$ then there exists a $i \in I$ such that $a \in U_i \Rightarrow [a, b] = \{a\} \subseteq \bigcup_{j \in \{i\}} U_j$ proving as $\{a\}$ is finite compactness

2. $(\boldsymbol{a < b})$ Here we proceed by contradiction and assume that $[a, b]$ is not compact. Then there exists by 12.71 and 12.249 a family $\{(a_i, b_i)\}_{i \in I}$ of open intervals such that $[a, b] \subseteq \bigcup_{i \in I} (a_i, b_i)$ that does not contains a finite cover. If we define now $\mathcal{A} = \{[x, y] | x, y \in \mathbb{R} \wedge a \leqslant x < y \leqslant b \text{ and } [x, y] \text{ is not covered by } a \text{ finite subcover of } \{(a_i, b_i)\}_{i \in I}\}$ and $\mathcal{M} = \{\{[x_i, y_i]\}_{i \in \{0, \ldots, n\}} | n \in \mathbb{N}_0 \wedge [x_i, y_i] \in \mathcal{A}\}$. Then if $\{[x_i, y_i]\}_{i \in \{0, \ldots, n\}} \in \mathcal{M}$ we have that $[x_n, y_n]$ is not covered by a finite subset of $\{[a_i, b_i]\}_{i \in I}$ and as $[x_n, y_n] \in \mathcal{A}$ we have that $a \leqslant x_n < y_n \leqslant b$, take now $c = x_n + \frac{y_n - x_n}{2}$. Then if $[x_i, c], [c, y_i]$ are both covered by finite sub-cover of $\{(a_i, b_i)\}_{i \in I}$ there exists finite $A, B \subseteq I$ such that $[x_n, c] \subseteq \bigcup_{i \in A} (a_i, b_i) \wedge [c, y_n] \subseteq \bigcup_{i \in B} (a_i, b_i) \Rightarrow [x_n, y_n] \subseteq \bigcup_{i \in A \bigcup B} (a_i, b_i)$ meaning that $[x_i, y_i]$ would be covered by a finite sub cover of $\{(a_i, b_i)\}_{i \in I}$ a contradiction. So either $[x_n, c]$ or $[c, y_n]$ is not covered by a finite sub cover of $\{(a_i, b_i)\}_{i \in I}$. Consider then the two cases to construct a $[x, y]$

   a. **($[\boldsymbol{x_n, c}]$ is not covered by $\boldsymbol{a}$ finite subcover of $\boldsymbol{\{(a_i, b_i)\}_{i \in I}}$)** take then $[x, y] = [x_n, c] \subseteq [x_n, y_n]$, $[x, y]$ is not covered by a finite sub cover of $\{(a_i, b_i)\}_{i \in I}$ and $y - x = x_n + \frac{y_n - x_n}{2} - x_n = \frac{y_n - x_n}{2}$ and $a \leqslant x_n \leqslant x < y \leqslant y_n \leqslant b$

   b. **($[\boldsymbol{x_n, c}]$ is covered by $\boldsymbol{a}$ finite subcover of $\boldsymbol{\{(a_i, b_i)\}_{i \in I}}$)** then $[c, y_n]$ is not covered by a finite subcover of $\{(a_i, b_i)\}_{i \in I}$. Take then $[x, y] = [c, y_n] \subseteq [x_n, y_n]$ and $y - x = y_n - \left(x_n + \frac{y_n - x_n}{2}\right) = (y_n - x_n) - \frac{y_n - x_n}{2} = \frac{y_n - x_n}{2}$ and $a \leqslant x_n \leqslant x < y < y_n \leqslant b$

   this defines a function $\delta: \mathcal{M} \rightarrow \mathcal{A}$ mapping $\{[x_i, y_i]\}_{i \in \{0, \ldots, n\}} \rightarrow \delta(\{[x_i, y_i]_{i \in \{0, \ldots, n\}}\}) = [x, y]$ such that $[x - y] \in \mathcal{A}$ and $[x, y] \subseteq [x_n, y_n]$. Using recursion and the fact that $[a, b] \in \mathcal{A}$ (see 5.24) we have found a sequence $\{[x_i, y_i]\}_{i \in \mathbb{N}_0}$ such that $\forall i \in \mathbb{N}_0$ we have that $[x_i, y_i]$ is not covered by a finite sub cover of $\{(a_i, b_i)\}_{i \in I}$, $a \leqslant x_i < y_i \leqslant b$, $[x_0, y_0] = [a, b]$ and $\forall i \in \mathbb{N}_0$ we have that $[x_{i+1}, y_{i+1}] \subseteq [x_i, y_i]$ and $y_i - x_i = \frac{b - a}{2^i}$. Using the previous theorem and the fact that $\left\{\frac{b - a}{2^i}\right\}_{i \in \mathbb{N}_0}$ converges to 0, we have $\exists c \in \mathbb{R}$ a such that $\{c\} = \bigcap_{i \in \mathbb{N}_0} [x_i, y_i] \subseteq [x_0, y_0] = [a, b]$. As $c \in [a, b]$ there exists a $i \in I$ such that $c \in (a_i, b_i) \Rightarrow a_i < c < b_i$. Take now $\varepsilon = \frac{1}{2} \cdot \min(c - a_i, b_i - c) > 0$ and choose a $j \in \mathbb{N}_0$ such that $y_j - x_j = |y_j - x_j| = \frac{b - a}{2^j} < \varepsilon$ $\left[\left\{\frac{b - a}{2^i}\right\}_{i \in \mathbb{N}_0}\right.$ converges to 0 (see 12.319)]. Then $c \in [x_j, y_j] \Rightarrow x_j \leqslant c \leqslant y_j \Rightarrow 0 \leqslant c - x_j \leqslant y_j - x_j < \varepsilon \leqslant c - a_i \Rightarrow -x_j < -a_i \Rightarrow a_i < x_j$ also $y_j - c < y_j - x_j < \varepsilon \leqslant b_i - c \Rightarrow y_j < b_i$ which proves that $a_i < x_j < y_j < b_i \Rightarrow [x_j, y_j] \subseteq (a_i, b_i)$ proving that $[x_j, y_j]$ is covered by a finite sub cover of $\{(a_i, b_i)\}_{i \in \mathbb{N}_0}$ which is a contradiction. So the assumption that $[a, b]$ is not compact leads to a contradiction and so $[a, b]$ is compact. $\square$

Theorem 12.258. *Let $\langle X, d_X\rangle$, $\langle Y, d_Y\rangle$ be metric spaces $K \subseteq X$ a compact subset and $f: X \rightarrow Y$ a continuous function then f is uniform continuous in K (see 12.150)*

Proof. Let $\varepsilon > 0$ then by continuity of f at $x \in K$ (see 12.147) there exists a $\delta_x \in \mathbb{R}_+$ such that if $d_X(x, y) < \delta_x$ then $d_Y(f(x), f(y)) < \frac{\varepsilon}{2}$. The family $\{B_{d_X}(x, \delta(x))\}_{x \in K}$ covers K so by compactness there exist a finite set $\{x_1, \ldots, x_n\}$ such that $K \subseteq \bigcup_{x \in \{x_1, \ldots x_n\}} B_{d_X}\left(x, \frac{\delta(x)}{2}\right)$. Take now $\delta = \min\left(\left\{\frac{\delta(x)}{2} | x \in \{x_1, \ldots, x_n\}\right\}\right)$. Then if $x, y \in K$ with $d_X(x, y) < \delta$ then there exists a $x_i \in \{x_1, \ldots, x_n\}$ such that $x \in B_{d_X}(x_i, \delta(x_i))$ so that $d_X(x_i, y) \leqslant d_X(x_i, x) + d_X(x, y) < \frac{\delta(x_i)}{2} + \delta \leqslant \frac{\delta(x_i)}{2} + \frac{\delta(x_i)}{2} = \delta(x_i)$, so we have that $d_Y(f(x), f(y)) \leqslant d(f(x), f(x_i)) + d(f(x_i), f(y)) \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ □

12.10.1 Filter bases in Topological spaces

To prove Tychonoff's theorem, that essentially says that any product of non empty compact topological spaces is compact in the product topology, we must develop a whole theoretical framework of filter bases.

Definition 12.259. *Let $\langle X, \mathcal{T}\rangle$ be a topological space. A filter base $\mathfrak{U}$ in X is a nonempty family $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$ of subsets of X ($\mathbb{U} \neq \emptyset$) such that*

1. *$\forall \alpha \in \mathbb{U}$ we have that $A_\alpha \neq \emptyset$*
2. *$\forall \alpha, \beta \in \mathbb{U}$ we have $\exists \gamma \in \mathbb{U}$ such that $A_\gamma \subseteq A_\alpha \bigcap A_\beta$ (this of course implies then that the intersection is non empty)*

Using mathematical induction it is then easy to prove the following theorem

Theorem 12.260. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$ a filter base then for each finite $\{\alpha_1, \ldots, \alpha_n\} \subseteq \mathbb{U}$ there exists a $\gamma \in \mathbb{U}$ such that $A_\gamma \subseteq \bigcap_{\alpha \in \{\alpha_1, \ldots, \alpha_n\}} A_\alpha$*

Proof. Let $S = \{n \in \mathbb{N} | \text{if } \{\alpha_1, \ldots, \alpha_n\} \subseteq \mathbb{U} \text{ then } \exists \gamma \in \mathbb{U} \text{ such that } A_\gamma \subseteq \bigcap_{\alpha \in \mathbb{U}} A_\alpha\}$ then we have

1. if $n = 1$ then if $\{a_1\} \in \mathbb{U}$ then $A_{\alpha_1} = \bigcup_{\alpha \in \{\alpha_1\}} A_\alpha$ so that $1 \in S$
2. if $n \in S$ then if $\{\alpha_1, \ldots, \alpha_{n+1}\} \subseteq \mathbb{U}$ then as $\{\alpha_1, \ldots, \alpha_n\} \subseteq \mathbb{U}$ and $n \in S$ there exists a $\gamma_1 \in \mathbb{U}$ such that $A_{\gamma_1} \in \bigcup_{\alpha \in \{\alpha_1, \ldots, \alpha_n\}} A_\alpha$, as $\mathfrak{U}$ is a filter base we can then find a $\gamma \in \mathbb{U}$ such that $A_\gamma \subset A_{\gamma_1} \bigcap A_{n+1} \subseteq \left(\bigcap_{\alpha \in \{\alpha_1, \ldots, \alpha_n\}} A_\alpha\right) \bigcap A_{n+1} = \bigcap_{\alpha \in \{\alpha_1, \ldots, a_{n+1}\}} A_\alpha$ and thus we have $n + 1 \in S$

Using mathematical induction we have then $S = \mathbb{N}$ proving the theorem. □

Definition 12.261. *Let $\langle X, \mathcal{T}\rangle$ be a topological space, $x \in X$ then $\mathfrak{U}(x) = \{U | U \in \mathcal{T} \wedge x \in U\}$ [the set of open sets containing x] is called the* ***neighborhood filter base*** *of x*

Theorem 12.262. *Let $\langle X, \mathcal{T}\rangle$ be a topological space, $x \in X$ then $\{U\}_{U \in \mathfrak{U}(x)}$ (see 2.101) is a filter-base, which is also noted as $\mathfrak{U}(x)$*

Proof.

1. If $U \in \mathfrak{U}(x)$ then $x \in U \Rightarrow U \neq \emptyset$

2. If $U, V \in \mathfrak{U}(x) \Rightarrow x \in U \bigcap V$ and as $U \bigcap V$ is open we have that $U \bigcap V \in \mathfrak{U}(x)$ so we found a $W = U \bigcap V \in \mathfrak{U}(x)$ such that $W = U \bigcap V \subseteq U \bigcap V$ □

Example 12.263. If $\langle X, \mathcal{T}\rangle$ and $\emptyset \neq A \subseteq X$ then $\{A\}_{A \in \{A\}}$ is trivially a filter base.

Theorem 12.264. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and let $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$ and $\mathfrak{W} = \{B_\beta\}_{\beta \in \mathbb{W}}$ then*

1. $\mathfrak{U} \bigcup \mathfrak{W} \underset{\text{defined}}{=} \{A_\alpha \bigcup B_\beta\}_{(\alpha,\beta) \in \mathbb{U} \times \mathbb{W}}$ *is a filter base*
2. *If $\forall \alpha \in \mathbb{U}, \forall \beta \in \mathbb{W}$ we have that $A_\alpha \bigcap B_\beta \neq \emptyset$ then $\mathfrak{U} \bigcap \mathfrak{W} \underset{\text{defined}}{=} \{A_\alpha \bigcap B_\beta\}_{(\alpha,\beta) \in \mathbb{U} \times \mathbb{W}}$* is a filter base

Proof.

1. **($\mathfrak{U} \bigcup \mathfrak{W}$ is *a* filter base)**
   a. $\forall (\alpha, \beta) \in \mathbb{U} \times \mathbb{W}$ we have $A_\alpha \neq \emptyset \wedge B_\beta \neq \emptyset \Rightarrow A_\alpha \bigcup B_\beta \neq \emptyset$
   b. $\forall (\alpha_1, \beta_1), (\alpha_2, \beta_2) \in \mathbb{U} \times \mathbb{W}$ there exists a $\gamma_1 \in \mathbb{U}, \gamma_2 \in \mathbb{W}$ such that $A_{\gamma_1} \subseteq A_{\alpha_1} \bigcap A_{\alpha_2} \wedge B_{\gamma_2} \subseteq B_{\beta_1} \bigcap B_{\beta_2}$ then $\gamma = (\gamma_1, \gamma_2) \in \mathbb{U} \times \mathbb{W}$ and $A_{\gamma_1} \bigcup B_{\gamma_2} \subseteq (A_{\alpha_1} \bigcap A_{\alpha_2}) \bigcup (B_{\beta_1} \bigcap B_{\beta_2}) \subseteq A_{\alpha_1} \bigcup B_{\beta_1} (A_{\alpha_2} \bigcup B_{\beta_2}) \Rightarrow A_{\gamma_1} \bigcup B_{\gamma_2} \subseteq (A_{\alpha_1} \bigcup B_{\beta_1}) \bigcap (A_{\alpha_2} \bigcap B_{\beta_2})$
2. **($\mathfrak{U} \bigcap \mathfrak{W}$ is *a* filter base (if $\forall (\alpha, \beta) \in \mathbb{U} \times \mathbb{W}$ we have that $A_\alpha \bigcap B_\beta \neq \emptyset$))**
   a. $\forall (\alpha, \beta) \in \mathbb{U} \times \mathbb{W}$ we have by the hypothesis that $A_\alpha \bigcap B_\beta \neq \emptyset$
   b. $\forall (\alpha_1, \beta_1), (\alpha_2, \beta_2) \in \mathbb{U} \times \mathbb{W}$ there exists a $\gamma_1 \in \mathbb{U}, \gamma_2 \in \mathbb{W}$ such that $A_{\gamma_1} \subseteq A_{\alpha_1} \bigcap A_{\alpha_2} \wedge B_{\gamma_2} \subseteq B_{\beta_1} \bigcap B_{\beta_2} \Rightarrow \gamma = (\gamma_1, \gamma_2) \subseteq \mathbb{U} \times \mathbb{W}$ and $A_{\gamma_1} \bigcap B_{\gamma_2} \subseteq (A_{\alpha_1} \bigcap A_{\alpha_2}) \bigcap (B_{\beta_1} \bigcap B_{\beta_2}) = (A_{\alpha_1} \bigcap B_{\beta_1}) \bigcap (A_{\alpha_2} \bigcap B_{\beta_2})$ □

Theorem 12.265. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$ be a filter base then it has the* ***finite intersection property*** *, that is $\forall \mathcal{B} \subseteq \mathbb{U}, \mathcal{B}$* is finite *we have $\bigcap_{\beta \in \mathcal{B}} A_\beta \neq \emptyset$*

Proof. Using 12.260 there exists a $\gamma \in \mathcal{B}$ such that $\emptyset \neq A_\gamma \subseteq \bigcap_{\beta \in \mathcal{B}} A_\beta$ □

Definition 12.266. *Let $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$ be a filter base in a topological space $\langle X, \mathcal{T}\rangle$ then we define*

1. *$\mathfrak{U}$ converges to $y \in X$ written as $\mathfrak{U} \rightarrow y$ iff $\forall U \in \mathcal{T}$ with $y \in U$ we have that $\exists \alpha \in \mathbb{U}$ with $A_\alpha \subseteq U$*
2. *$\mathfrak{U}$ accumulates at $y \in X$ (or y is a accumulation point of $\mathfrak{U}$) written as $\mathfrak{U} \succ y$ iff $\forall U \in \mathcal{T}$ with $y \in U$ we have $\forall \alpha \in \mathbb{U}$ that $U \bigcap A_\alpha \neq \emptyset$*

Remark 12.267. If $\mathfrak{U} \succ y$ then $\forall \alpha \in \mathbb{U}$ we have $[\forall U \in \mathcal{T} \vdash y \in U$ we have $U \bigcap A_\alpha \neq \emptyset] \underset{12.19}{\Leftrightarrow} [y \in \overline{A_a}]$ so $\mathfrak{U} \succ y$ is equivalent with $y \in \bigcap_{\alpha \in \mathbb{U}} \overline{A_\alpha}$

Example 12.268. Let $\langle X, \mathcal{T}\rangle$ be a topological space then $\forall x \in X$ we have $\mathfrak{U}(x) \rightarrow x$

Proof. If $U \in \mathcal{T}$ and $x \in U$ then by definition $U \in \mathfrak{U}(x)$ so that we found a $U \in \mathfrak{U}(x)$ with $U \subseteq U$ □

Definition 12.269. *Let $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}, \mathfrak{W} = \{B_\beta\}_{\beta \in \mathbb{W}}$ be two filter bases on a topological space $\langle X, \mathcal{T}\rangle$ then $\mathfrak{W}$ is subordinate to $\mathfrak{U}$ noted by $\mathfrak{W} \gg \mathfrak{U}$ (or $\mathfrak{U} \ll \mathfrak{W}$) iff $\forall \alpha \in \mathbb{U}$ we have that $\exists \beta \in \mathbb{W}$ such that $B_\beta \subseteq A_\alpha$*

Definition 12.270. *Let $\mathfrak{U}=\{A_\alpha\}_{\alpha\in\mathbb{U}}, \mathfrak{W}=\{B_\beta\}_{\beta\in\mathbb{W}}$ be two filter bases on a topological space $\langle X,\mathcal{T}\rangle$ then $\mathfrak{U}\subseteq\mathfrak{W}$ iff $\forall\alpha\in\mathbb{U}$ there exists a $\beta\in\mathbb{W}$ such that $A_\alpha=B_\beta$*

Theorem 12.271. *Let $\mathfrak{U}=\{A_\alpha\}_{\alpha\in\mathbb{U}}, \mathfrak{W}=\{B_\beta\}_{\beta\in\mathbb{W}}$ be two filter bases on a topological space $\langle X,\mathcal{T}\rangle$ then the following hold:*

1. *If $\mathfrak{U}\subseteq\mathfrak{W}$ then $\mathfrak{W}\gg\mathfrak{U}$*
2. *If $\mathfrak{W}\gg\mathfrak{U}$ then $\forall\beta\in\mathbb{W}$ we have that $\forall\alpha\in\mathfrak{U}$ that $A_\alpha\bigcap B_\beta\neq\emptyset$*
3. *$\mathfrak{U}\rightarrow x$ if and only if $\mathfrak{U}\gg\mathfrak{U}(x)$*
4. *If $\mathfrak{W}=\{A_\alpha\}_{\alpha\in\mathbb{W}}, \mathfrak{W}'=\{A'_\alpha\}_{\alpha\in\mathbb{W}'}, \mathfrak{W}''=\{A''_a\}_{a\in\mathbb{W}''}$ are filter bases such that $\mathfrak{W}\gg\mathfrak{W}'$ and $\mathfrak{W}'\gg\mathfrak{W}''$ then $\mathfrak{W}\gg\mathfrak{W}''$*

Proof.

1. Let $\alpha\in\mathbb{U}$ then as $\mathfrak{U}\subseteq\mathfrak{W}$ there exists a $b\in\mathbb{W}$ such that $A_\alpha=B_\beta$ so that $B_\alpha\subseteq A_\alpha$ so that by definition we have $\mathfrak{W}\gg\mathfrak{U}$
2. This is proved by contradiction, so assume that $\exists\alpha\in\mathbb{U}$ and $\exists\beta\in\mathbb{W}$ such that $A_\alpha\bigcap B_\beta=\emptyset$ then since $\mathfrak{W}\gg\mathfrak{U}$ $\exists\gamma\in\mathbb{W}$ such that $B_\gamma\subseteq A_\alpha$ and then there exists a $\delta\in\mathbb{W}$ such that $B_\delta\subseteq B_\gamma\bigcap B_\beta\subseteq A_\alpha\bigcap B_\beta$ which as $B_\delta\neq\emptyset$ means that $A_\alpha\bigcap B_\beta\neq\emptyset$ a contradiction.
3. $\mathfrak{U}\rightarrow x\Leftrightarrow\forall U$ open with $x\in U$ we have that $\exists\alpha\in\mathbb{U}$ such that $A_\alpha\subseteq U\Leftrightarrow$ $\forall U\in\mathfrak{U}(x)$ we have $\exists\alpha\in\mathbb{U}$ such that $A_\alpha\subseteq U\Leftrightarrow\mathfrak{U}\gg\mathfrak{U}(x)$
4. Let $\alpha''\in\mathbb{W}''$ then because $\mathfrak{W}'\gg\mathfrak{W}''$ there exists a $\alpha'\in\mathbb{W}'$ such that $A'_{\alpha'}\subseteq A''_{\alpha''}$, because $\mathfrak{W}\gg\mathfrak{W}'$ there exists a $\alpha\in\mathbb{W}$ such that $A_\alpha\subseteq A'_{\alpha'}\subseteq A''_{\alpha''}$ and thus $\mathfrak{W}\gg\mathfrak{W}''$ □

Theorem 12.272. *Let $\langle X,\mathcal{T}\rangle$ be a topological space then X is Hausdorff $\Leftrightarrow$ each filter base converges to exactly one point.*

Proof.

1. **(⇒)** Assume that X is Hausdorff, $\mathfrak{U}=\{A_\alpha\}_{\alpha\in\mathbb{U}}$ so that $\mathfrak{U}\rightarrow x$ and $\mathfrak{U}\rightarrow y$ and $x\neq y$ then there exists U,V open with $x\in U\wedge y\in V\wedge U\bigcap V=\emptyset$. From $\mathfrak{U}\rightarrow x$ and $\mathfrak{U}\rightarrow y$ there exists $\alpha,\beta\in\mathbb{U}$ such that $A_\alpha\subseteq U\wedge A_\beta\subseteq V\Rightarrow$ $\emptyset\neq A_\alpha\bigcap A_\beta\subseteq U\bigcap V=\emptyset$ a contradiction. So we must have that $x=y$
2. **(⇐)** Assume that X is not Hausdorff then $\exists x,y\in Xx\neq y$ such that $\forall U$, V open with $x\in U\wedge y\in V$ we have $U\bigcap V\neq\emptyset$ then by 12.264 we have that $\mathfrak{W}=\mathfrak{U}(x)\bigcap\mathfrak{U}(y)$ is a filter base. If $x\in U(U$ open$)\Rightarrow U\in\mathfrak{U}(x)$ and as $X\in\mathfrak{U}(y)$ then $U=U\bigcap X\in\mathfrak{W}$ and thus $\mathfrak{W}\rightarrow x$, if $y\in V(V$ open$)\Rightarrow$ $V\in\mathfrak{U}(y)$ and as $X\in\mathfrak{U}(x)$ we have that $V=X\bigcap V\in\mathfrak{W}$ proving that $\mathfrak{W}\rightarrow y$. By the hypothesis we have then that $x=y$ contradicting $x\neq y$. So X must be Hausdorff. □

Theorem 12.273. *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $\mathfrak{U}=\{A_\alpha\}_{\alpha\in\mathbb{U}}$, $\mathfrak{W}=\{B_\beta\}_{\beta\in\mathbb{W}}$ be filter bases then*

1. *If $\mathfrak{U}\rightarrow x$ then $\mathfrak{U}\succ x$ and if X is also* Hausdorff *then if $\mathfrak{U}\succ y\Rightarrow x=y$*

2. *Let* $\mathfrak{W} \gg \mathfrak{U}$ *then*

    a. $\mathfrak{U} \rightarrow x \Rightarrow \mathfrak{W} \rightarrow x$

    b. $\mathfrak{W} \succ x \Rightarrow \mathfrak{U} \succ x$

Proof.

1. *Let* U *be open with* $x \in U$ *then* $\exists \alpha \in \mathbb{U}$ *such that* $A_\alpha \subseteq U$. *Given* $\beta \in \mathbb{U}$ *then as (see 12.265)* $\emptyset \neq A_\alpha \bigcap A_\beta \subseteq U \bigcap A_\beta \Rightarrow \mathfrak{U} \succ x$. *Now if* X *is Hausdorff and let* $\mathfrak{U} \succ y$ *with* $x \neq y$ *then* $\exists U, V$ *open such that* $x \in U, y \in V$ *and* $U \bigcap V = \emptyset$ *then* $\exists \alpha \in \mathbb{U}$ *such that* $A_\alpha \subseteq U$ *(as* $\mathfrak{U} \rightarrow x$*) so that* $A_\alpha \bigcap V \subseteq U \bigcap V = \emptyset$ *contradicting* $\mathfrak{U} \succ y$ *so we must have* $x = y$.

2. *Assume that* $\mathfrak{W} \gg \mathfrak{U}$ *then*

    a. *Assume that* $\mathfrak{U} \rightarrow x$ *then if* U *is open with* $x \in U$ *then* $\exists \alpha \in \mathbb{U}$ *such that* $A_\alpha \subseteq U$ *and as* $\mathfrak{W} \gg \mathfrak{U}$ *then* $\exists \beta \in \mathbb{W}$ *such that* $A_\beta \subseteq A_\alpha \subseteq U$ *and thus* $\mathfrak{W} \rightarrow x$

    b. *Assume that* $\mathfrak{W} \succ x$ *then if* U *is open such that* $x \in U$ *then* $\forall \beta \in \mathbb{W}$ *we have* $U \bigcap B_\beta = \emptyset$, *now if* $\alpha \in \mathbb{U}$ *then as* $\mathfrak{W} \gg \mathfrak{U}$ *there* $\exists \gamma \in \mathbb{W}$ *such that* $B_\gamma \subseteq A_\alpha \Rightarrow \emptyset \neq U \bigcap B_\gamma \subseteq U \bigcap A_\alpha \Rightarrow \mathfrak{U} \succ x$ □

Theorem 12.274. *Let* $\langle X, T \rangle$ *be a topological space and let* $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$ *be a filter-base then*

1. $\mathfrak{U} \rightarrow x$ *if and only if* $\forall \mathfrak{W} = \{B_\beta\}_{\beta \in \mathbb{W}}$ *(a filter base) with* $\mathfrak{W} \gg \mathfrak{U}$ *we have that there exists a filter-base* $\mathfrak{C} = \{C_\gamma\}_{\gamma \in \mathbb{C}}$ *with* $\mathfrak{C} \gg \mathfrak{W}$ *with* $\mathfrak{C} \rightarrow x$

2. $\mathfrak{U} \succ x$ *if and only if* $\exists \mathfrak{W} = \{B_\beta\}_{\beta \in \mathbb{B}}$ *(a filter base) with* $\mathfrak{W} \gg \mathfrak{U}$ *such that* $\mathfrak{W} \rightarrow x$

Proof.

1.

    a. **(⇒)** Assume $\mathfrak{U} \rightarrow x$ then if $\mathfrak{W} \gg \mathfrak{U} \underset{12.273}{\Rightarrow} \mathfrak{W} \rightarrow x$ and as trivially we have $\mathfrak{W} \gg \mathfrak{W}$ we can use $\mathfrak{C} = \mathfrak{W}$

    b. **(⇐)** Assume that $\mathfrak{U} \nrightarrow x$ ($\mathfrak{U}$ does not converge to x) then $\exists U$ open with $x \in U$ such that $\forall \alpha \in \mathbb{U}$ we have $A_\alpha \nsubseteq U \Rightarrow A_\alpha \bigcap (X \setminus U) \neq \emptyset$ then $\mathfrak{W} = \{A_\alpha \bigcap (X \setminus U)\}_{\alpha \in \mathbb{U}}$ forms a filter basis (see 12.263 and 12.264) and we have trivially $\mathfrak{W} \gg \mathfrak{U}$. Then by the hypothesis $\exists \mathfrak{C} = \{C_\gamma\}_{\gamma \in \mathbb{C}}$ such that $\mathfrak{C} \gg \mathfrak{W}$ and $\mathfrak{C} \rightarrow x \underset{12.273}{\Rightarrow} \mathfrak{W} \succ x$ which means that $\forall \alpha \in \mathbb{U}$ we have that $(A_\alpha \bigcap (X \setminus U)) \bigcap U \neq 0$ which is a contradiction because $(A_\alpha \bigcap (X \setminus U)) \bigcap U \subseteq (X \setminus U) \bigcap U = \emptyset$. So we must have that $\mathfrak{U} \rightarrow x$

2.

    a. **(⇒)** Assume that $\mathfrak{U} \succ x$ then if U is open with $x \in U$ then $\forall \alpha \in \mathbb{U}$ we have that $A_\alpha \bigcap U \neq \emptyset$, so using 12.261 and 12.264 we have then that $\mathfrak{W} = \mathfrak{U} \bigcap \mathfrak{U}(x) = \{A_\alpha \bigcap U\}_{(\alpha, U) \in \mathbb{U} \times \mathfrak{U}(x)}$ is a filter base. As $\forall \alpha \in \mathbb{U}$ we have that $A_\alpha \bigcap U \subseteq A_a$ so that $\mathfrak{W} \gg \mathfrak{U}$. As by the definition of a filter base $\mathbb{U} \neq \emptyset$ there exists a $\alpha \in \mathbb{U}$ and then we have that $A_\alpha \bigcap U \subseteq U$ proving that $\mathfrak{W} \rightarrow x$

b. ($\Leftarrow$) If there exists a $\mathfrak{W} \gg \mathfrak{U}$ with $\mathfrak{W} \rightarrow x$ then using 12.273 we have $\mathfrak{W} \succ x$ and using 12.273 again we have then $\mathfrak{U} \succ x$ □

Definition 12.275. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and $A \subseteq X$ then a filter base $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$ is said to be* ***on*** *A if and only if $\forall \alpha \in \mathbb{U}$ we have $A_\alpha \subseteq A$*

Theorem 12.276. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and let $A \subseteq X$ then we have the following equivalence $y \in \bar{A} \Leftrightarrow \exists \mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$ which is a filter base on A with $\mathfrak{U} \rightarrow x$*

Proof.

1. ($\Rightarrow$) If $y \in \bar{A}$ then $\forall U$ open with $y \in U$ we have $A \bigcap U \neq \emptyset$, using 12.263,12.261 and 12.264 we have that $\mathfrak{W} = \{A\} \bigcap \mathfrak{U}(x) = \{A \bigcap U\}_{U \in \mathfrak{U}(x)}$ is a filter base, which is on A [as $A \bigcap U \subseteq A$]. If U is open with $x \in U$ then $U \in \mathfrak{U}(x)$ and $A \bigcap U \subseteq U$ proving that $\mathfrak{W} \rightarrow x$
2. ($\Leftarrow$) If $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$ is a filter base on A with $\mathfrak{U} \rightarrow x$ then if U is open and $x \in U$ open then there exists a $\alpha \in \mathbb{U}$ such that $A_\alpha \subseteq U$ and as $A_\alpha \neq \emptyset$ and $A_\alpha \subseteq A$ we have that $\emptyset \neq A_\alpha = A_\alpha \bigcap A \subseteq U \bigcap A \Rightarrow x \in \bar{A}$ □

Theorem 12.277. *Let $\langle X, \mathcal{T}\rangle$, $\langle Y, \boldsymbol{S}\rangle$ be topological spaces and let $f: X \rightarrow Y$ be a function. If $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$ is a filter base in X then $f(\mathfrak{U}) \underset{\text{define}}{=} \{f(A_\alpha)\}_{\alpha \in \mathbb{U}}$ is a filter base in Y*

Proof. First as $\mathbb{U} \neq \emptyset$ and $\forall \alpha \in \mathbb{U}$ we have $A_a \neq \emptyset \Rightarrow f(A_\alpha) \neq \emptyset \Rightarrow \{f(A_\alpha)\}_{\alpha \in \mathbb{U}}$ is a non empty family of non empty sets. Second if $\alpha, \beta \in \mathbb{U}$ then there exists a $\gamma \in \mathbb{U}$ such that $A_\gamma \subseteq A_\alpha \bigcap A_\beta \Rightarrow f(A_\gamma) \subseteq f(A_\alpha \bigcap A_\beta) \subseteq f(A_\alpha) \bigcap f(A_\beta)$ so that $\{f(A_\alpha)\}_{\alpha \in \mathbb{U}}$ is a filter base. □

Theorem 12.278. *Let $\langle X, \mathcal{T}\rangle$, $\langle Y, \boldsymbol{S}\rangle$ be topological spaces and let $f: X \rightarrow Y$ be a function then f is continue at x if and only if $f(\mathfrak{U}(x)) \rightarrow f(x)$*

Proof.

1. ($\Rightarrow$) Assume that f is continuous at x and let V open with $f(x) \in V$ then $\exists U$ open with $x \in U$ such that $f(U) \subseteq V \Rightarrow f(\mathfrak{U}(x)) \rightarrow f(x)$
2. ($\Leftarrow$) Let V be open with $f(x) \in V$ then as $f(\mathfrak{U}(x)) \rightarrow f(x)$ and $\mathfrak{U}(x) = \{U\}_{U \in \mathfrak{U}(x)}$ there exists a $U \in \mathfrak{U}(x) \Rightarrow x \in U, U$ open such that $f(U) \subseteq V \Rightarrow f$ is continuous at x □

Theorem 12.279. *Let $\langle X, \mathcal{T}\rangle$, $\langle Y, \boldsymbol{S}\rangle$ be topological spaces and let $f: X \rightarrow Y$ be a function then f is continue on X if and only if $\forall x \in X$ and for every filter base $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$ in X with $\mathfrak{U} \rightarrow x$ we have $f(\mathfrak{U}) \rightarrow f(x)$*

Proof.

1. ($\Rightarrow$) Assume that f is continuous and let $\mathfrak{U} \rightarrow x$ then by 12.271 we have $\mathfrak{U} \gg \mathfrak{U}(x)$ so if U is open and $x \in U$ then $\exists \alpha \in \mathbb{U}$ such that $A_\alpha \subseteq U \Rightarrow f(A_\alpha) \subseteq f(U)$ proving that $f(\mathfrak{U}) \gg f(\mathfrak{U}(x))$. As f is continuous we have by 12.278 that $f(\mathfrak{U}(x)) \rightarrow f(x)$ so using 12.273 and $f(\mathfrak{U}) \gg f(\mathfrak{U}(x))$ we have $f(\mathfrak{U}) \rightarrow f(x)$

2. ($\Leftarrow$) Let $A \subseteq X$ and let $y \in f(\bar{A})$ then $\exists x \in \bar{A}$ such that $y = f(x)$ so that by 12.276 there exists a filter base $\mathfrak{U} = \{A_\alpha\}_{\alpha \langle A \rangle}$ on A so that $\mathfrak{U} \rightarrow x$. As $\mathfrak{U}$ is a filter base on A we have $\forall a \in \mathbb{U}$ that $A_\alpha \subseteq A$ so that $f(A_\alpha) \subseteq f(A)$ and thus $f(\mathfrak{U})$ is on $f(A)$. This last together with $f(\mathfrak{U}) \rightarrow f(x)$ gives by 12.276 that $f(x) \in \overline{f(A)}$ or $f(\bar{A}) \subseteq \overline{f(A)}$ and thus by 12.138 that f is continuous. □

Theorem 12.280. *Let $\{\langle X_i, \mathcal{T}_i \rangle\}_{i \in I}$ be a family of topological spaces then a filter-base $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$ in $X = \prod_{i \in I} X_i$ converges to $y \in \prod_{i \in I} X_i$ (using the product topology) if and only if $\forall i \in I$ we have that $\pi_i(\mathfrak{U}) \rightarrow \pi_i(y) = y(i) = y_i$*

Proof.

1. ($\Rightarrow$) Let $\mathfrak{U} \rightarrow y$ then as π_i is continuous (if we use the product topology) we have by 12.279 that $\pi_i(\mathfrak{U}) \rightarrow \pi_i(y)$
2. ($\Leftarrow$) Let $y \in X = \prod_{i \in I} X_i$ and U open in X with $y \in U$ then there exists a B in the basis of the product topology such that $y \in B \subseteq U$. By the definition of the basis of a product topology there exists a finite set $\mathcal{A} \subseteq I$ such that $B = \bigcap_{i \in \mathcal{A}} \pi_i^{-1}(U_i)$ where U_i is a open set in X_i (see 12.38). So $\forall i \in \mathcal{A}$ we have that $\pi_i(y) \in U_i$. As we have by the hypothesis that $\forall i \in I$ we have that $\pi_i(\mathfrak{U}) \rightarrow \pi_i(y)$ there exists $\forall i \in \mathcal{A}$ a $\alpha_i \in \mathbb{U}$ such that $\pi_i(A_{\alpha_i}) \subseteq U_i \Rightarrow A_{\alpha_i} \subseteq \pi_i^{-1}(U_i) \Rightarrow \bigcap_{i \in \mathcal{A}} A_{\alpha_i} \subseteq \bigcap_{i \in \mathcal{A}} \pi_i^{-1}(U_i) = B$. Using 12.260 and the finiteness of $\mathcal{A}$ there exists a $\gamma \in \mathbb{U}$ such that $A_\gamma \subseteq \bigcap_{i \in \mathcal{A}} A_{\alpha_i} \subseteq B \subseteq U \Rightarrow \mathfrak{U} \rightarrow y$ □

Definition 12.281. *A filter base $\mathfrak{U}$ in a topological space $\langle X, \mathcal{T} \rangle$ is called **maximal** or a **ultra filter** if $\forall \mathfrak{W}$, $\mathfrak{W}$ a filter base in X with $\mathfrak{W} \gg \mathfrak{U}$ we have that $\mathfrak{U} \gg \mathfrak{W}$.*

Theorem 12.282. *A filter base $\mathfrak{U}$ in a topological space $\langle X, \mathcal{T} \rangle$ is maximal if and only if $\forall A \subseteq X$ one and only one of the two sets A and $X \setminus A$ contains a member of $\mathfrak{U}$.*

Proof.

1. ($\Rightarrow$) Assume that $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$ is maximal and let $A \subseteq X$ then if there exists a $\alpha, \beta \in \mathbb{U}$ such that $A_\alpha \subseteq A$ and $A_\beta \subseteq X \setminus A$ then as by definition of a filter base we have $A_\alpha \bigcap A_\beta \neq \emptyset$ we have that $\emptyset \neq A_\alpha \bigcap A_\beta \subseteq A_\alpha \bigcap X \setminus A = \emptyset$ a contradiction, so we can not have a $\alpha, \beta \in \mathbb{U}$ such that $A_\alpha \subseteq A \wedge A_\beta \subseteq X \setminus A$. Assume now that $\forall \alpha \in \mathbb{U}$ we have that $A_\alpha \subseteq A$ then we have that $\forall \alpha \in \mathbb{U}$ that $(X \setminus A) \bigcap A_\alpha \neq \emptyset$ so $\mathfrak{W} = \{(X \setminus A) \bigcap A_\alpha\}_{\alpha \in \mathbb{U}}$ forms a filter base (see 12.263 and 12.264) and we have $\mathfrak{W} \gg \mathfrak{U}$ [if $\alpha \in \mathbb{U}$ then $(X \setminus A) \bigcap A_\alpha \subseteq A_\alpha$] so that by maximality we must have that $\mathfrak{U} \gg \mathfrak{W}$. So $\forall \alpha \in \mathbb{U}$ there exists a $\beta \in \mathbb{U}$ so that $A_\beta \subseteq (X \setminus A) \bigcap A_\beta \Rightarrow A_\beta \subseteq X \setminus A$
2. ($\Leftarrow$) Let $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$ be a filter base such that $\forall A \subseteq X$ one and only one of the two sets $A, X \setminus A$ contains a member of $\mathfrak{U}$. Let then $\mathfrak{W} = \{B_\beta\}_{\beta \in \mathbb{W}}$ such that $\mathfrak{W} \gg \mathfrak{U}$. If $\beta \in \mathbb{W}$ then [by taking $B_\beta = A$ in the above) we have the existence of a $\alpha \in A$ such that either $A_\alpha \subseteq B_\beta$ or $A_\alpha \subseteq X \setminus B_\beta$. The last case would mean that $A_\alpha \bigcap B_\beta = \emptyset$ contradicting 12.271 so that we are left with $A_\alpha \subseteq B_\beta$ proving that $\mathfrak{U} \gg \mathfrak{W}$ and thus maximality. □

Theorem 12.283. *Let $\mathfrak{U}=\{A_\alpha\}_{\alpha\in\mathbb{U}}$ be any filter base in a topological space $\langle X, \mathcal{T}\rangle$ then there exists a maximal filter base $\mathfrak{W}$ in X with $\mathfrak{W}\gg\mathfrak{U}$*

Proof. Let $\mathcal{A}=\{\mathfrak{W}|\mathfrak{W}\text{ is }a\text{ filterbase such that }\mathfrak{W}\gg\mathfrak{U}\}$ then we have that $\mathcal{A}\neq\emptyset$ because $\mathfrak{U}\gg\mathfrak{U}$ and thus $\mathfrak{U}\in\mathcal{A}$. We define now a pre-order (see 2.131) $\leqslant$ on $\mathcal{A}$ defined by $\mathfrak{W}\leqslant\mathfrak{W}'\Leftrightarrow\mathfrak{W}'\gg\mathfrak{W}$ this is a pre-order because

1. **(reflexivity)** $\mathfrak{W}\gg\mathfrak{W}\Rightarrow\mathfrak{W}\leqslant\mathfrak{W}$
2. **(transitivity)** Assume that $\mathfrak{W}\leqslant\mathfrak{W}'\wedge\mathfrak{W}'\leqslant\mathfrak{W}''\Rightarrow\mathfrak{W}'\gg\mathfrak{W}\wedge\mathfrak{W}''\gg\mathfrak{W}'\underset{12.271}{\Rightarrow}\mathfrak{W}''\gg\mathfrak{W}\Rightarrow\mathfrak{W}\leqslant\mathfrak{W}''$

Take now $\mathcal{C}\subseteq\mathcal{A}$ be any chain in the pre-ordered set $\langle\mathcal{A},\leqslant\rangle$ and prove that it has a upper bound in $\mathcal{A}$. We have two cases to consider:

1. **($\mathcal{C}=\emptyset$)** Then we have vacuous that $\forall\mathfrak{C}\in\mathcal{C}$ we have that $\mathfrak{C}\leqslant\mathfrak{U}$ so $\mathfrak{U}\in\mathfrak{A}$ is a upper bound of $\mathcal{C}$
2. **($\mathcal{C}\neq\emptyset$)** Define $\mathfrak{B}=\{A|\exists\mathfrak{C}=\{C_\gamma\}_{\gamma\in\mathbb{C}}\in\mathcal{C}\vdash\exists\gamma\in\mathbb{C}\text{ with }A=C_\gamma\}$ then $\mathfrak{B}=\{B\}_{B\in\mathfrak{B}}$ is a filter base because
   a. If $B\in\mathfrak{B}$ then $\exists\mathfrak{C}=\{C_\gamma\}_{\gamma\in\mathbb{C}}\in\mathcal{C}$ such that $\exists\gamma\in\mathbb{C}$ with $B=C_\gamma\neq\emptyset$
   b. If $B_1,B_2\in\mathfrak{B}$ then there exists $\mathfrak{C}_1=\{C^1_\gamma\}_{\gamma\in\mathbb{C}_1}$, $\mathfrak{C}_2=\{C^2_\gamma\}_{\gamma\in\mathbb{C}_2}$, $\mathfrak{C}_1$, $\mathfrak{C}_2\in\mathcal{C}$ such that $\exists\gamma_1\in\mathbb{C}_1\wedge\exists\gamma_2\in\mathbb{C}_2$ such that $B_1=C^1_{\gamma_1}\wedge B_2=C^2_{\gamma_2}$. Now as $\mathcal{C}$ is a chain we have either:
      i. **($\mathfrak{C}_1\leqslant\mathfrak{C}_2$)** $\Rightarrow\mathfrak{C}_2\gg\mathfrak{C}_1\Rightarrow\exists\gamma\in\mathbb{C}_2$ such that $C^2_\gamma\subseteq C^1_{\gamma_1}=B_1$ and because $\mathfrak{C}_2$ is a filter base there exists a $\beta\in\mathbb{C}_2$ such that $C^2_\beta\subseteq C^2_\gamma\bigcap C^2_{\gamma_2}\subseteq B_1\bigcap B_2$ so that we have found a $B_3=C^2_\beta\in\mathfrak{B}$ such that $B_3\subseteq B_1\bigcap B_2$
      ii. **($\mathfrak{C}_2\leqslant\mathfrak{C}_1$)** $\Rightarrow\mathfrak{C}_1\gg\mathfrak{C}_2\Rightarrow\exists\gamma\in\mathbb{C}_1$ such that $C^1_\gamma\subseteq C^2_{\gamma_2}=B_2$ and because $\mathfrak{C}_1$ is a filter base there exists a $\beta\in\mathbb{C}_1$ such that $C^1_\beta\subseteq C^1_\gamma\bigcap C^1_{\gamma_2}\subseteq B_2\bigcap B_1=B_1\bigcap B_2$ so that we have found a $B_3=C^1_\beta\in\mathfrak{B}$ such that $B_3\subseteq B_1\bigcap B_2$

   Using (a) and (b) we have proved that $\mathfrak{B}$ is a filter base. Now as $\mathcal{C}\neq\emptyset$ there exists a $\mathfrak{C}=\{C_\gamma\}_{\gamma\in\mathbb{C}}\in\mathcal{C}\subseteq\mathcal{A}$ so that $\mathfrak{C}\gg\mathfrak{U}$ and thus if $\alpha\in\mathbb{U}$ we have a $\gamma\in\mathbb{C}$ such that $C_\gamma\subseteq A_\alpha$ and as by definition we have that $C_\gamma\in\mathfrak{B}$ it follows that $\mathfrak{B}\gg\mathfrak{U}$ so that $\mathfrak{B}\in\mathcal{A}$. Finally $\forall\mathfrak{C}=\{C_\gamma\}_{\gamma\in\mathbb{C}}\in\mathcal{C}$ we have that $\forall\gamma\in\mathbb{C}$ that $C_\gamma\subseteq C_\gamma$ and because $C_\gamma\in\mathfrak{B}$ (by definition of $\mathfrak{B}$) we have that $\mathfrak{B}\gg\mathfrak{C}$ or $\mathfrak{C}\leqslant\mathfrak{B}$. This completes the proof that $\mathfrak{B}$ is a upper bound in $\mathcal{A}$ of $\mathcal{C}$. Using 2.226 there exists a maximum element $\mathfrak{M}$ of $\mathcal{A}$ in the pre-order on $\mathcal{A}$. This means (see 2.160) that for every $\mathfrak{A}\in\mathcal{A}$ with $\mathfrak{A}\gg\mathfrak{M}\Rightarrow\mathfrak{M}\leqslant\mathfrak{A}$ we have $\mathfrak{A}\leqslant\mathfrak{M}\Rightarrow\mathfrak{M}\gg\mathfrak{A}$. Also if $\mathfrak{Q}$ is a filter base in X with $\mathfrak{Q}\gg\mathfrak{M}$ then as $\mathfrak{M}\gg\mathfrak{U}$ we have that $\mathfrak{Q}\gg\mathfrak{U}$ so that $\mathfrak{Q}\in\mathcal{A}$ and thus by the previous that $\mathfrak{M}\gg\mathfrak{Q}$ proving that $\mathfrak{M}$ is a maximal filter base (or ultra filter) with $\mathfrak{M}\gg\mathfrak{U}$. □

Theorem 12.284. *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $\mathfrak{U}=\{A_\alpha\}_{\alpha\in\mathbb{U}}$ a maximal filter base then $\mathfrak{U}\succ x\Leftrightarrow\mathfrak{U}\rightarrow x$ for every $x\in X$*

Proof. Because of 12.273 we must only prove that $\mathfrak{U}\succ x\Rightarrow\mathfrak{U}\rightarrow x$ for a maximal filter base. So let $x\in U$, U open then by 12.282 there exists a $\alpha\in\mathbb{U}$ such that either $A_\alpha\subseteq U$ or $A_\alpha\subseteq X\setminus U$. Now as $\mathfrak{U}\succ x$ we must have that $A_\alpha\bigcap U\neq\emptyset$ contradicting $A_\alpha\subseteq X\setminus U$ so we must have that $A_\alpha\subseteq U$ or $\mathfrak{U}\rightarrow x$ □

Theorem 12.285. *Let $\langle X, \mathcal{T}_X\rangle$, $\langle Y, \mathcal{T}_Y\rangle$ be topological spaces and let $\mathfrak{U} = \{A_\alpha\}_{\alpha\in\mathbb{U}}$ is a maximal filter base in X then for any function $f': X \rightarrow Y$ we have that $f(\mathfrak{U})$ is a maximal filter base in Y*

Proof. By 12.277 we have that $f(\mathfrak{U}) = \{f(A_\alpha)\}_{\alpha\in\mathbb{U}}$ is a filter base. To prove maximality we use 12.282, so let $A \subseteq Y$ and take $f^{-1}(A)$ and $X \setminus f^{-1}(A)$ then by maximality of $\mathfrak{U}$ there exists a $\alpha \in \mathbb{U}$ such that $A_\alpha \subseteq f^{-1}(A)$ or $A_\alpha \subseteq X \setminus f^{-1}(A)$ so that we have either $f(A_a) \subseteq f(f^{-1}(A)) = A$ or $f(A_\alpha) \subseteq f(X \setminus f^{-1}(A)) \subseteq Y \setminus A$ proving that $f(\mathfrak{U})$ is maximal. □

All this hard work is to prove Tychonoffs theorem on the product of compact spaces. The following theorem creates the relation between compactness and filter bases.

Theorem 12.286. *Let $\langle X, \mathcal{T}\rangle$ be a topological space then the following are equivalent*

1. *X is compact*
2. *For each family $\{A_\alpha\}_{\alpha\in\mathcal{A}}$ of closed sets in X with $\bigcap_{\alpha\in\mathcal{A}} A_\alpha = \emptyset$ we have the existence of a finite sub-family $\{A_\beta\}_{\beta\in\mathcal{B}}$ [where $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\} \subseteq \mathcal{A}$] such that $\bigcap_{\beta\in\mathcal{B}} A_\beta = \emptyset$ (or in other words if $\{A_\alpha\}_{\alpha\in\mathcal{A}}$ is a family such that for every finite $\mathcal{B} \subseteq \mathcal{A}$ we have that $\bigcap_{\alpha\in\mathcal{B}} A_\alpha \neq \emptyset$ (the finite intersection property) then we must have $\bigcap_{\alpha\in\mathcal{A}} A_\alpha \neq \emptyset$)*
3. *Each filter base $\{A_\alpha\}_{\alpha\in\mathbb{U}}$ in X has at least one accumulation point*
4. *Each maximal filter base in X converges*

Proof.

1. **(1 ⇒ 2)** Let $\{A_\alpha\}_{\alpha\in\mathcal{A}}$ be a family of closed sets with $\bigcap_{\alpha\in\mathcal{A}} A_\alpha = \emptyset$ then $X = X \setminus \emptyset = X \setminus (\bigcap_{\alpha\in\mathcal{A}} A_\alpha) = \bigcup_{\alpha\in\mathcal{A}} (X \setminus A_\alpha)$ and because $X \setminus A_\alpha$ is open $\forall \alpha \in \mathcal{A}$ we have by compactness the existence of a finite $\mathcal{B} \subseteq \mathcal{A}$ such that $X = \bigcup_{\beta\in\mathcal{B}} (X \setminus A_\beta) \Rightarrow \emptyset = X \setminus X = X \setminus (\bigcup_{\beta\in\mathcal{B}} (X \setminus A_\beta)) = \bigcap_{\beta\in\mathcal{B}} (X \setminus (X \setminus A_\beta)) = \bigcap_{\beta\in\mathcal{B}} A_\beta$
2. **(2 ⇒ 1)** Let $\{U_\alpha\}_{\alpha\in\mathcal{A}}$ be a family of open sets such that $X = \bigcup_{\alpha\in\mathcal{A}} U_\alpha \Rightarrow \emptyset = X \setminus X = X \setminus (\bigcup_{\alpha\in\mathcal{A}} U_\alpha) = \bigcap_{\alpha\in A} (X \setminus U_\alpha)$ and as we have that $\forall \alpha \in \mathcal{A}$ that $X \setminus U_\alpha$ is closed we have by the hypothesis that $\exists \mathcal{B} \subseteq \mathcal{A}$, $\mathcal{B}$ finite such that $\emptyset = \bigcap_\beta (X \setminus U_\beta) \Rightarrow X \setminus \emptyset = X \setminus (\bigcap_{\beta\in\mathcal{B}} (X \setminus U_\beta)) = \bigcup_{\beta\in\mathcal{B}} (X \setminus (X \setminus U_\beta)) = \bigcup_{\beta\in\mathcal{B}} U_\beta$ proving compactness
3. **(2 ⇒ 3)** Let $\mathfrak{U} = \{A_\alpha\}_{\alpha\in\mathbb{U}}$ be a filter base, then $\forall \mathbb{B} \subseteq \mathbb{U} \vdash \mathbb{B}$ is finite we have that $\emptyset \neq \bigcap_{\beta\in\mathbb{B}} A_\beta$ (see 12.260) then $\emptyset \neq \bigcap_{\beta\in\mathbb{B}} A_\beta \subseteq \bigcap_{\beta\in\mathbb{B}} \overline{A_\beta}$ and then by (2) we have $\emptyset \neq \bigcap_{\alpha\in\mathbb{A}} \overline{A_\alpha} \Rightarrow \exists x \in \bigcap_{\alpha\in\mathbb{A}} \overline{A_\alpha}$ and then by 12.267 we have that x is a accumulation point of $\mathfrak{U}$.
4. **(3 ⇒ 4)** Let $\mathfrak{U}$ be a maximal filter base then by (3) there exists a $x \in X$ such that $\mathfrak{U} \succ x$, and then using 12.284 we have $\mathfrak{U} \rightarrow x$
5. **(4 ⇒ 3)** Let $\mathfrak{U}$ be a filter base then using 12.283 there exist a maximal filter base $\mathfrak{W} \gg \mathfrak{U}$ and then by (4) we have then $\exists x \in X$ such that $\mathfrak{W} \rightarrow x \underset{12.284}{\Rightarrow} \mathfrak{W} \succ x$ and then using 12.273 we have $\mathfrak{U} \succ x$

6. **(3 ⇒ 2)** Let $\{F_\alpha\}_{\alpha\in\mathcal{A}}$ be a family of closed sets such that $\forall\mathcal{B}\subseteq\mathcal{A},\mathcal{B}$ finite we have $\bigcap_{\beta\in\mathcal{B}}F_\beta\neq\emptyset$ then if $\mathfrak{U}=\{\bigcap_{\beta\in\mathcal{B}}A_\beta|\mathcal{B}\subseteq\mathcal{A},\mathcal{B}\text{ is finite}\}$ we form the family $\mathfrak{U}=\{A\}_{A\in\mathfrak{U}}$. We prove now that $\mathfrak{U}$ is a filter base, first of all we have by the hypothesis that $\bigcap_{\beta\in\mathcal{B}}F_\beta\neq\emptyset$, also if $B_1,B_2\in\mathfrak{U}$ then $\exists\mathcal{B}_1,\mathcal{B}_2\subseteq\mathcal{A}$, $\mathcal{B}_1,\mathcal{B}_2$ finite and thus $\mathcal{B}_1\bigcup\mathcal{B}_2$ is finite such that $B_1\bigcap B_2=(\bigcap_{\beta\in\mathcal{B}_1}F_\beta)\bigcap(\bigcap_{\beta\in\mathcal{B}_2}F_\beta)=\bigcap_{\beta\in\mathcal{B}_1\bigcup\mathcal{B}_2}F_\beta\in\mathfrak{U}$ [[if $x\in(\bigcap_{\beta\in\mathcal{B}_1}F_\beta)\bigcap(\bigcap_{\beta\in\mathcal{B}_2}F_\beta)$ then $\forall\beta\in\mathcal{B}_1$, $\forall\alpha\in\mathcal{B}_1$ we have $x\in F_\beta\wedge x\in F_\alpha\Rightarrow$ if $\gamma\in\mathcal{B}_1\bigcup\mathcal{B}_2$ then $x\in F_\gamma$, if $x\in\bigcap_{\gamma\in\mathcal{B}_1\bigcup\mathcal{B}_2}F_\gamma$ then if $\gamma\in\mathcal{B}_1\bigcup\mathcal{B}_2$ we have $x\in F_\gamma$ so that if $\alpha\in\mathcal{B}_1$ and $\beta\in\mathcal{B}_1$ then $x\in(\bigcap_{\beta\in\mathcal{B}_1}F_\beta)\bigcap(\bigcap_{\beta\in\mathcal{B}_2}F_\beta)$]] so by (3) $\exists x\in X$ such that $\mathfrak{U}\succ x$ or using 12.267 we have $x\in\bigcap_{B\in\mathfrak{U}}\bar{B}$ so $\forall B\in\mathfrak{U}$ we have $x\in\bar{B}$ now as $\forall\alpha\in\mathcal{A}$ we have $\{\alpha\}$ is finite and thus $F_\alpha=\bigcap_{\beta\in\{\alpha\}}F_\alpha\in\mathfrak{U}\Rightarrow x\in\overline{F_\alpha}\underset{A_\alpha\,\text{is closed}}{=}F_\alpha\Rightarrow x\in\bigcap_{\alpha\in\mathcal{A}}F_\alpha\Rightarrow\bigcap_{\alpha\in\mathcal{A}}F_\alpha\neq\emptyset$ □

12.10.2 Product of compact spaces

Let's use now filter bases to prove Tychonoff's theorem

Theorem 12.287. (Tychonoff's) *Let $\{\langle X_i,\mathcal{T}_i\rangle\}_{i\in I}$ be a family non empty topological spaces then $\prod_{i\in I}X_i$ is compact in the product topology if and only if $\forall i\in I$ we have that X_i is compact*

Proof.

1. **(⇒)** As the projection map $\pi_i:\prod_{j\in I}X_j\rightarrow X_i$ is continuous (see 12.144) we have using 12.243 that $X_i=\pi_i(\prod_{j\in I}X_j)$ is compact.
2. **(⇐)** Assume that $\forall i\in I$ we have that X_i is compact. Let $\mathfrak{U}$ be a maximum filter base in $\Pi_{i\in I}X_i$ then using 12.277 we have that $\forall j\in I$ that $\pi_j(\mathfrak{U})$ is a maximum filter base. Using the compactness of X_i we have using 12.286 that $\exists x_i\in X_i$ such that $\pi_i(\mathfrak{U})\rightarrow x_i$. Define then $x\in\prod_{j\in I}X_j$ by $x(i)=x_i$ so that $\pi_i(\mathfrak{U})\rightarrow\pi_i(x)$. Using 12.280 we have then that $\mathfrak{U}\rightarrow x$ which by 12.286 means that $\prod_{i\in I}X_i$ is convergent. □

Corollary 12.288. *Let $\{\langle X_i,\mathcal{T}_i\rangle\}_{i\in I}$ be a family of non-empty topological spaces then if $\forall i\in I$ we have that $C_i\subseteq X_i$ then $\prod_{i\in I}C_i$ is compact if and only if $\forall i\in I$ we have that C_i is compact (using the subspace topology).*

Proof. First note that the subspace topology of $\prod_{i\in I}C_i$ is equal to the product topology of the subspace topologies (see 12.40). As by definition a subset is compact is if it is compact in the subspace topology we can use the previous theorem to prove this theorem. □

Example 12.289. Let $\langle\mathbb{R}^n,\mathcal{T}\rangle$ be the set of real tupple's together with the product topology (based on the trivial normed topology on $\langle\mathbb{R},||\rangle$ then $\prod_{i\in\{1,\ldots,n\}}[a_i,b_i]$ is compact.

Proof. Using 12.257 we have $\forall i\in\{1,\ldots,n\}$ that $[a_i,b_i]$ is a compact set in R (using the norm topology generated in $\mathbb{R}$, so using the previous corollary to Tychonoff's theorem we have that $\prod_{i\in\{1,\ldots,n\}}[a_i,b_i]$ is compact. □

Theorem 12.290. (Heine-Borel) *Let $\langle \mathbb{R}^n, \mathcal{T} \rangle$ be the set of real tupple's together with the product topology (based on the trivial normed topology on $\langle \mathbb{R}, || \rangle$ then a subset of $C \subseteq \mathbb{R}^n$ is compact if and only if C is bounded and closed.*

Proof.

1. **(⇒)** Using 12.242 and 12.67 we have that C is bounded. As by 12.219 and 12.67 we have that $\langle \mathbb{R}^n, \mathcal{T} \rangle$ is Hausdorff and using 12.244 that C is closed.
2. **(⇐)** Assume now that C is closed and bounded. If C is empty then it is trivially compact so for the rest of the proof we assume that C is not empty. Thus there exists a $c \in C$. If we use the canonical norm $\|x\| = \max(|\pi_i(x)||i \in \{1, \ldots, n\})$ and the metric $d(x, y) = \|x - y\| = \max(\{|\pi_i(x - y)||i \in \{1, \ldots, n\}\}) = \max(\{|\pi_i(x) - \pi_i(y)||i \in \{1, \ldots, n\}\})$ then we have as C is bounded that there exists a $M > 0$ such that $\forall y \in C$ we have that $d(c, y) \leqslant M$ or $\forall i \in \{1, \ldots, n\}$ we have that $|\pi_i(c) - \pi_i(y)| \leqslant M \Rightarrow \pi_i(y) \in [\pi_i(c) - M, \pi_i(c) + M]$ and thus $C \subseteq \prod_{i \in \{1, \ldots, n\}} [\pi_i(c) - M, \pi_i(c) + M]$ which is compact by 12.257 and the corollary to Tychonoff's theorem. As $C = C \bigcap \left(\prod_{i \in \{1, \ldots, n\}} [\pi_i(c) - M, \pi_i(c) + M] \right)$ is closed in the subspace topology of the compact set $\prod_{i \in \{1, \ldots, n\}} [\pi_i(c) - M, \pi_i(c) + M]$ (as C is closed) we can use 12.245 to prove that C is compact. □

Theorem 12.291. *Let $\langle X, \mathcal{T} \rangle$ be a topological space and $f: X \to \mathbb{R}$ be a continuous real function (using the canonical topology on $\mathbb{R}$). Then for every compact $C \subseteq X$ there exists a $M_C > 0$ such that $\forall x \in C$ we have $|f(x)| \leqslant M_C$*

Proof. First if $C = \emptyset$ then the theorem is vacuously satisfied so assume that $C \neq \emptyset$. Using 12.243 it follows that $f(C)$ is compact so that by Heine-Borel (see 12.290) we find that $f(C)$ is bounded and closed. So $\exists k > 0$ so that $\forall y_1, y_2 \in f(C)$ we have that $|y_1 - y_2| \leqslant k$ and thus if $x_1, x_2 \in C \Rightarrow f(x_1), f(x_2) \in f(C) \Rightarrow |f(x_1) - f(x_2)| \leqslant k$. As $C \neq \emptyset$ there exists a $c_0 \in C$, take then $M_C = k + |f(c_0)|$ then $\forall c \in C$ we have $|f(c)| \leqslant |f(c) - f(c_0)| + |f(c_0)| \leqslant k + |f(c)| = M_C$. □

Theorem 12.292. *Let $\langle X, \mathcal{T} \rangle$ be a topological space and $f: X \to \mathbb{R}$ be a continuous function. Then for every non empty compact subset $C \subseteq X$ there exists a $x_1, x_2 \in C$ such that $\forall c \in C$ we have that $f(x_1) \leqslant f(c) \leqslant f(x_2)$.*

Proof. By 12.243 we have that $f(C)$ is compact so that by 12.290 $f(C)$ is bounded and closed. Using then 12.291 there exists a M_C such that $\forall c \in C$ we have $|f(c)| \leqslant M_C \Rightarrow -M_C \leqslant f(c) \leqslant M_C$. By the lower (upper) bound property of the real's (see 9.43 and 2.176) and the fact that $f(C) \neq \emptyset$ there exists a $M_1 = \inf(f(C))$ and $M_2 = \sup(f(C))$. If now $U \subseteq \mathbb{R}$ is a open set with $M_1 \in U$ (or $M_2 \in U$) then there exists a open ball $]a, b[$ such that $M_1 \in]a, b[\subseteq U \Rightarrow a < M_1 < b$ (or $M_2 \in]a, b[\subseteq U \Rightarrow a < M_2 < b$) and by the definition of inf (or sup) we have the existence of a $y \in f(C)$ such that $a < M_1 \leqslant y < b \Rightarrow y \in U$ (or $a < y \leqslant M_2 < b \Rightarrow y \in U$) proving that $y \in f(C) \bigcap U \Rightarrow f(C) \bigcap U \neq \emptyset$. Using 12.20 we have then that M_1, $M_2 \in \overline{f(C)} \underset{f(C) \text{ is compact} \Rightarrow f(C) \text{ is closed}}{=} f(C)$ so that there exists a $c_1, c_2 \in C$ such that $M_1 = f(c_1)$, $M_2 = f(c_2)$ and taking in account the definition of inf and sup we have then that $\forall c \in C$ we have $f(c_1) = M_1 = \inf(f(C)) \leqslant f(c) \leqslant \sup(f(C)) = M_2 = f(c_2) \Rightarrow f(c_1) \leqslant f(c) \leqslant f(c_2)$. □

Theorem 12.293. *All norms on $\mathbb{R}^n$ are equivalent (see 12.82)*

Proof. Assume that $\|\|$ is the maximum norm in $\mathbb{R}^n$ and $\|\|^*$ another norm in $\mathbb{R}^n$ then by 12.159 we have that $\|\|^*: \mathbb{R}^n \rightarrow \mathbb{R}$ is a continuous mapping (using the maximum norm topology on $\mathbb{R}^n$). Define then $S^{n-1} = \{x \in \mathbb{R}^n | \|x\| = 1\}$ which is closed because $\{1\}$ is closed [as a metric space is Hausdorff (see 12.219) and every finite set in a Hausdorff space is closed (see 12.220)] and $S^{n-1} = \|\|^{-1}(\{1\})$ together with 12.138. S^{n-1} is also bound because if $x, y \in S^{n-1}$ then $\|x - y\| \leqslant \|x\| + \|y\| \leqslant 1 + 1 = 2$. According to the Heine Borell theorem (see 12.290) and the fact that the product topology on $\mathbb{R}^n$ is the topology generated by the maximum norm (see 12.79) we have then that S^{-1} is compact in $\mathbb{R}^n$ using the maximum norm. Using the previous theorem (see 12.292) there exists a $x_1, x_2 \in S^{n-1}$ such that $\forall x \in S^{-1}$ we have $\|x_1\|^* \leqslant \|x\|^* \leqslant \|x_2\|^* \Rightarrow \alpha \leqslant \|x\|^* \leqslant \beta$ if we define $\alpha = \|x_1\|^* > 0$ ($x_1 \neq 0$ as $\|x_1\| = 1$) and $\beta = \|x_2\|^* > 0$ ($x_2 \neq 0$ as $\|x_2\| = 1$). If now $x \in \mathbb{R}^n$ then we have either

1. $(\boldsymbol{x} = \mathbf{0})$ $\Rightarrow \|x\| = 0 = \|x\|^* \Rightarrow \alpha \cdot \|x\| \leqslant \|x\|^* \leqslant \beta \cdot \|x\|$
2. $(\boldsymbol{x} \neq \mathbf{0})$ then $\|x\| \neq 0$ and $\left\|\frac{1}{\|x\|} \cdot x\right\| = \frac{\|x\|}{\|x\|} = 1 \Rightarrow \alpha \leqslant \left\|\frac{1}{\|x\|} \cdot x\right\|^* \leqslant \beta \Rightarrow \alpha \leqslant \frac{1}{\|x\|} \cdot x \leqslant \beta \Rightarrow \alpha \cdot \|x\| \leqslant \|x\|^* \leqslant \beta \cdot \|x\|$

Using 12.82 we find that $\|\|$ and $\|\|^*$ are equivalent, so every norm generates the same topology as $\|\|$ and thus all the norms are equivalent. □

Theorem 12.294. *Let $\langle X, \|\|\rangle$ be a normed space, $n \in \mathbb{N}$ such that $\varphi: \mathbb{R}^n \rightarrow X$ a isomorphism then if $C \subseteq X$ that C is compact if and only if C is closed and bounded*

Proof. Define $\|\|_n: \mathbb{R}^n \rightarrow X$ by $\|x\|_n = \|\varphi(x)\|$ then we have

1. If $x \in \mathbb{R}^n$ then $\|x\|_n = \|\varphi(x)\| \geqslant 0$
2. If $\alpha \in \mathbb{R}$, $x \in \mathbb{R}^n$ then $\|\alpha \cdot x\|_n = \|\varphi(\alpha \cdot x)\| \underset{\varphi \text{ is linear}}{=} \|\alpha \cdot \varphi(x)\| \underset{\|\| \text{ is a norm}}{=} |\alpha| \cdot \|\varphi(x)\| = |\alpha| \cdot \|x\|_n$
3. If $x, y \in \mathbb{R}^n$ then $\|a + y\|_n = \|\varphi(x + y)\| \underset{\varphi \text{ is linear}}{=} \|\varphi(x) + \varphi(y)\| \leqslant \|\varphi(x)\| + \|\varphi(y)\| = \|x\|_n + \|y\|_n$
4. If $\|x\|_n = 0$ then $\|\varphi(x)\| = 0 \underset{\|\| \text{ is a norm}}{\Rightarrow} \varphi(x) \underset{\varphi \text{ is a isomorphism}}{\Rightarrow} x = 0$

The above proves that $\|\|_n$ is a norm an by construction a isometry, using 12.167 we have then that φ is a homeomorphism and isometry between the normed spaces $\langle \mathbb{R}^n, \|\|_n\rangle$ and $\langle X, \|\|\rangle$. Let now $C \subseteq X$ be a closed set then as φ is a homeomorphism we have that $(\varphi^{-1})(C)$ is closed. Further as C is bounded we have that there exists a $0 < M$ such that $\forall x, y \in C$ we have $\|x - y\| \leqslant M$. Hence if $x', y' \in (\varphi^{-1})(C)$ we have that $x' = (\varphi^{-1})(x), y' = (\varphi^{-1})(y)$ where $x, y \in C$ so that $x = \varphi(x'), y = \varphi(y')$ hence $\|x' - y'\|_2 = \|\varphi(x' - y')\| = \|\varphi(x') - \varphi(y')\| = \|x - y\| \leqslant M$ proving that $(\varphi^{-1})(C)$ is closed and bounded in the topology generated by $\|\|_n$ on $\mathbb{R}^n$. Using the previous theorem (12.293) we have that, if $\mathcal{T}$ is the product topology generated by the maximum norm (see 12.79), that:

If $C \subseteq X$ is closed and bounded in $\langle X, \|\|\rangle$ then $(\varphi^{-1})(C)$ is closed and bounded in $\langle \mathbb{R}^n, \mathcal{T}\rangle$ (12.25)

Using the Heine Borell theorem (see 12.290) we have then that $(\varphi^{-1})(C)$ is compact and as φ is a homeomorphism and thus continuous we have by 12.243 that $C = \varphi((\varphi^{-1})(C))$ is compact. Hence we conclude that

$$\text{If } C \subseteq X \text{ is closed and bounded in } \langle X, \|\|\rangle \text{ then } C \text{ is compact in } \langle X, \|\|\rangle \tag{12.26}$$

If C is compact then as $\varphi: \mathbb{R}^n \to X$ is a homemorphism between $\langle \mathbb{R}_n, \|\|_n\rangle$ and $\langle X, \|\|\rangle$ we have that $\varphi^{-1}: X \to \mathbb{R}^n$ is continuous so that by 12.243 we have that $(\varphi^{-1})(C)$ is compact. As by the previous theorem (12.293) we have that the topology generated by $\|\|_n$ is the same as the product topology generated by the maximum norm (see 12.79) if follows from the Heine Borel theorem (see 12.290) we have that $(\varphi^{-1})(C)$ is closed and bounded. As φ^{-1} is continuous we have $C = \varphi((\varphi^{-1})(C)) = (\varphi^{-1})^{-1}((\varphi^{-1})(C))$ is closed. Further as $\varphi^{-1}(C)$ is bounded there exists $M > 0$ such that $\forall x, y \in \varphi^{-1}(C)$ we have $\|x - y\|_n < M$, hence if $x, y \in C$ then $\varphi^{-1}(x), \varphi^{-1}(y) \in (\varphi^{-1})(C) \Rightarrow \|\varphi^{-1}(x) - \varphi^{-1}(y)\|_n \leqslant M$. So $\|x - y\| = \|\varphi(\varphi^{-1}(x)) - \varphi(\varphi^{-1}(y))\| = \|\varphi(\varphi^{-1}(x) - \varphi^{-1}(y))\| = \|\varphi^{-1}(x) - \varphi^{-1}(y)\|_n \leqslant M$ proving that

$$\text{If } C \subseteq X \text{ is compact in } \langle X, \|\|\rangle \text{ then } C \text{ is closed and bounded in } \langle X, \|\|\rangle \tag{12.27}$$ □

We can use this to prove that closed bounded sets in $\mathbb{C}$ are compact

Corollary 12.295. *Let $\langle \mathbb{C}, \|\|\rangle$ be the complex space with a norm then if $C \subseteq \mathbb{C}$ we have that C is compact if and only if C is closed and bounded*

Proof. Using 10.180 we have that there exists a isomorphism $\mathcal{C}: \mathbb{R}^2 \to \mathbb{C}$ between the vector spaces $\langle \mathbb{R}^2, +, \cdot\rangle$ and $\langle \mathbb{C}, +, \cdot\rangle$. The theorem is then prove by applying the previous theorem (see 12.294). □

Theorem 12.296. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be finite dimensional spaces then every linear function $L: X \to Y$ is continuous.*

Proof. Using 12.169 there exists $\varphi_1: X \to \mathbb{R}^n$, $\varphi_2: Y \to \mathbb{R}^m$ and norms $\|\|_1, \|\|_2$ on $\mathbb{R}^n$ and $\mathbb{R}^m$ where φ_1, φ_2 are isometries, homeomorphism's and isomorphism's (using the norms $\|\|_1, \|\|_2$). So given a linear map $L: X \to Y$ we have that $L' = \varphi_2 \circ L \circ \varphi_1^{-1}$: $\mathbb{R}^n \to \mathbb{R}^m$ is linear (as isomorphism's are linear). By 12.176 the linear mapping $L': \mathbb{R}^n \to \mathbb{R}^m$ is continuous if we use the maximum norms on $\mathbb{R}^n, \mathbb{R}^m$, by the equivalence of the norms of $\mathbb{R}^n$, $\mathbb{R}^m$ (see 12.293) we have also that L' is continuous if we use the norms $\|\|_1, \|\|_2$. Then as $L = \varphi_2^{-1} \circ L' \circ \varphi_1$ we have that L is continuous as φ_2^{-1}, L' and φ_1 are continuous. □

Theorem 12.297. *Every finite n-dimensional normed space $\langle X, \|\|_X\rangle$ is isomorphic and homeomorphic with every other n-dimensional normed space $\langle Y, \|\|_Y\rangle$.*

Proof. Using 12.169 there exists $\varphi_1: X \to \mathbb{R}^n$, $\varphi_2: Y \to \mathbb{R}^n$ and norms $\|\|_1, \|\|_2$ on $\mathbb{R}^n$ and $\mathbb{R}^n$ where φ_1, φ_2 are isometries, homeomorphism's and isomorphism's (using the norms $\|\|_1, \|\|_2$). Using 12.293 we have that $\|\|_1$ and $\|\|_2$ must be equivalent so that φ_1, φ_2 are also homeomorphism's using the topology generated by $\|\|_1$ so that $\varphi_2^{-1} \circ \varphi_1: X \to Y$ is a homeomorphism as a composition of homeomorphism (see 12.164 and 12.163) and also a isomorphism (by 10.182 and 10.181). □

12.11 Convergence

12.11.1 Sequencess

Notation 12.298. *Given a $m \in \mathbb{N}_0$ we note by $\{m, \ldots, \infty\}$ the set $\{n \in \mathbb{N}_0 | n \geqslant m\}$. For example $\{0, \ldots, \infty\} = \mathbb{N}_0$ and $\{1, \ldots, \infty\} = \mathbb{N}$.*

Definition 12.299. *Let X be a set and $k \in \mathbb{N}_0$ then $\{x_i\}_{i \in \{k, \ldots, \infty\}} \subseteq X$ is called a s sequence (here we extend the concept of sequences as families of the form $\{x_i\}_{i \in \mathbb{N}_0}$)*

Definition 12.300. *Let $\langle X, \leqslant \rangle$ be a partially ordered set and $\{x_i\}_{i \in \{k, \ldots, \infty\}} \subseteq X$ a sequence such that $\forall i \in \{k, \ldots, \infty\}$ we have $x_i \leqslant x_{i+1}$ [or $x_{i+1} \leqslant x_i$] then $\{x_i\}_{i \in \{k, \ldots, \infty\}}$ is called a **increasing (decreasing) seqeunce**.*

The following lemma will be used a lot of times when dealing with increasing (decreasing) sequences.

Lemma 12.301. *Let $\langle X, \leqslant \rangle$ be a partially ordered set and $\{x_i\}_{i \in \{k, \ldots, \infty\}}$ a increasing (decreasing) sequence then $\forall n, m \in \{k, \ldots, \infty\}$* with *$n \leqslant m$ we have $x_n \leqslant x_m$ ($x_m \leqslant x_n$)*

Proof. This is proved by induction so let $n \in \{k, \ldots., \infty\}$ and take $\mathcal{S}_n = \{i \in \mathbb{N}_0 | x_n \leqslant x_{n+i}\}$ then we have

$0 \in \mathcal{S}_n$. as $x_n = x_{n+0} \leqslant x_{n+0}$

$i \in \mathcal{S} \Rightarrow i + 1 \in \mathcal{S}_n$. as $i \in \mathcal{S}_n$ we have $x_n \leqslant x_{n+i}$ and as $x_{n+i} \leqslant x_{(n+i)+1}$ we have $x_n \leqslant x_{n+(i+1)}$ proving that $i + \in \mathcal{S}_n$

Using mathematical induction we have then that $\mathcal{S}_n = \mathbb{N}_0$. If now $n \leqslant m$ then $m = n + (m - n)$ and as $(m - n) \in \mathbb{N}_0 = \mathcal{S}_n$ we have $x_n \geqslant x_{n+(m-n)} = x_m$ □

Definition 12.302. *If $\langle X, \mathcal{T} \rangle$ is a topological space then a sequence $\{x_i\}_{i \in \{m, \ldots, \infty\}}$ in X is convergent to $x \in X$ if $\forall U \in \mathcal{T}$ with $x \in U$ there exists a $N \in \{m, \ldots, \infty\}$ such that $\forall n \in \mathbb{N}_0$ with $n \geqslant N$ we have $x_n \in U$. We call x a limit of the sequence $\{x_i\}_{i \in \{m, \ldots, \infty\}}$.*

Theorem 12.303. *If $\langle X, \mathcal{T} \rangle$ is a Hausdorff topological space then a convergent sequence $\{x_i\}_{i \in \{m, \ldots, \infty\}}$ in X has a unique limit. We note this limit as $\lim_{i \to \infty} x_i$*

Proof. Suppose that x, y are two limits of $\{x_i\}_{i \in \{m, \ldots, n\}}$ with $x \neq y$ then by the Hausdorff property there exists open sets U, V with $x \in U \wedge y \in V \wedge U \bigcap V = \emptyset$. By the definition of convergence there exists $N_x, N_y \in \{m, \ldots, \infty\}$ such that if $n \geqslant N_x$ then $x_n \in U$ and if $k \geqslant N_y$ then $x_n \in V$ so if $n \geqslant \max(N_x, N_y)$ then $x_n \in U \wedge x_n \in V \Rightarrow x_n \in U \bigcap V = \emptyset$ a contradiction, so we must have that $x = y$. □

The following theorem shows that convergence is not dependent on where a sequence start.

Theorem 12.304. *If $\langle X, \mathcal{T}\rangle$ is a topological space and $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ a sequence in X then*

1. *If $\{x_i\}_{i\in\{m,\ldots,n\}}$ converges to x we have that $\forall n\geqslant m$ that the subsequence $\{x_i\}_{i\in\{n,\ldots,\infty\}}$ converges to x*
2. *If there exists a $m\geqslant n$ such that $\{x_i\}_{i\in\{n,\ldots,\infty\}}$ converges to x then $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ converges to x*

Proof.

1. Given a U open containing x there exists by the hypothese a $N'\in\{m,\ldots,\infty\}$ such that $\forall i\geqslant N'$ we have $x_i\in U$. Take then $N=\max(n,N')$ then $N\in\{n,\ldots\infty\}$ and $i\geqslant N'$ such that $x_i\in U$ proving that $\{x_i\}_{i\in\{n,\ldots,\infty\}}$ converges to x.
2. If $\{x_i\}_{i\in\{n,\ldots,\infty\}}$ converges to x where $n\geqslant m$ then there exists a $N\in\{n,\ldots,\infty\}\Rightarrow N\in\{m,\ldots\infty\}$ such that if $i\geqslant N$ we have $x_i\in U$ proving that $\{x_i\}_{i\in\{m,\ldots\infty\}}$ converges to x> □

For metric spaces we define a limit of a sequence as follows

Definition 12.305. *If $\langle X,d\rangle$ is a pseudo-metric space then a sequence $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ in X is convergent to $x\in X$ if $\forall\varepsilon\in\mathbb{R}_+$ there exists a $N\in\{m,\ldots,\infty\}$ such that if $n\geqslant N$ then $d(x,x_n)<\varepsilon$. x is called a limit of $\{x_i\}_{i\in\{m,\ldots,\infty\}}$.*

The following theorem proves that both definitions are the same.

Theorem 12.306. *If $\langle X,d\rangle$ is a pseudo-metric space and $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ is a sequence then this sequence is convergent to x in the topological space sense if and only if it is convergent in the pseudo-metric sense to x.*

Proof.

1. **(⇒)** Let x be the limit of $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ in the topological space sense then if $\varepsilon\in\mathbb{R}_+$ we have for the open set $B_d(x,\varepsilon)$ that $x\in B_d(x,\varepsilon)$ and there exists a $N\in\{m,\ldots\infty\}$ such that if $n\geqslant N$ we have $x_n\in B_d(x,\varepsilon)\Rightarrow d(x,x_n)<\varepsilon$ proving that x is a limit of $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ in the pseudo-metric sense.
2. **(⇐)** Let x be the limit of $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ in the metric space then if U is a open set such that $x\in U$ then by the definition of the metric topology there exists a $\varepsilon\in\mathbb{R}_+$ such that $x\in B_d(x,\varepsilon)\subseteq U$. Using the convergence in the pseudo-metric sense there exists a $N\in\{m,\ldots,\infty\}$ such that if $n\geqslant N$ then we have $d(x,x_n)<\varepsilon\Rightarrow x_n\in B_d(x,\varepsilon)\subseteq U\Rightarrow x_n\in U$. Proving that x is the limit of $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ in the topological sense. □

Note that the limits of a sequence do not have to be unique, however the following theorem proves that in the case of a metric space the limit is unique.

Theorem 12.307. *If $\langle X,d\rangle$ is a metric space then if the sequence $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ is convergent it has a unique limit*

Proof. This is trivial as $\langle X,d\rangle$ is a Hausdorff (see 12.219) and Hausdorff spaces has a unique limit (see 12.303). □

Definition 12.308. *If $\langle X, d\rangle$ is a metric space and $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ is convergent then we note the unique limit as* $\lim_{i\rightarrow\infty} x_i$.

Definition 12.309. *If $\langle X, d\rangle$ is a pseudo-metric space then $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ is a Cauchy sequence if $\forall\varepsilon\in\mathbb{R}_+$ there exists a $N\in\mathbb{N}$ such that $\forall n,k\geqslant N$ we have $d(x_n,x_k)<\varepsilon$*

Theorem 12.310. *If $\langle X, d\rangle$ is a pseudo-metric space then $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ is a Cauchy sequence iff $\forall\varepsilon\in\mathbb{R}_+$ there exists a $N\in\{m,\ldots,\infty\}$ such that $\forall n\in\mathbb{N}_0$ we have that $d(x_N,x_{N+n})<\varepsilon$*

Proof.

1. **($\Rightarrow$)** If $\{x_i\}_{i\in\mathbb{N}}$ is Cauchy then given $\varepsilon\in\mathbb{R}_+$ there exist a $N\in\{m,\ldots,\infty\}$ such that for $n,k\geqslant N$ we have $d(x_n,x_k)<\varepsilon$ so if $n\in\mathbb{N}_0$ then $N\geqslant N$ and $N+n\geqslant N$ and thus we have $d(x_N,x_{N+n})$

2. **($\Leftarrow$)** Let $\varepsilon\in\mathbb{R}_+$ and find then a $N\in\{m,\ldots,\infty\}$ such that if $n\in\mathbb{N}_0$ then $d(x_N,x_{N+n})<\frac{\varepsilon}{2}$. If now $n,k\geqslant N$ then $n-N,k-N\in\mathbb{N}_0$ so that $d(x_n, x_k)=d(x_{(n-N)+N},x_{(k-N)+N})\leqslant d(x_{(n-N)+N},x_N)+d(x_N,x_{(k-N)+N})<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$ □

Theorem 12.311. *If $\langle X, d\rangle$ is a pseudo-metric space and $\{x_i\}_{i\in m,\ldots,\infty}$ is convergent then it is Cauchy.*

Proof. If $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ is convergent to x then if $\varepsilon\in\mathbb{R}_+$ there exists a $N\in\{m,\ldots,\infty\}$ such that if $n\geqslant N$ that $d(x,x_n)<\frac{\varepsilon}{2}$. So if $n,m\geqslant N$ then $d(x_n,x_m)\leqslant d(x_n, x)+d(x,x_m)<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$ proving that it is Cauchy. □

We have defined the convergence in a metric space, lets create equivalent definitions in a normed space.

Definition 12.312. *If $\langle X, \|\|\rangle$ is a pseudo-normed space then a sequence $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ of elements of X is convergent to $x\in X$ if $\forall\varepsilon\in\mathbb{R}_+$ we have that $\exists N\in\{m,\ldots,\infty\}$ so that $\forall n$ with $n\geqslant N$ we have $\|x-x_n\|<\varepsilon$*

Example 12.313. If $\langle X, \|\|\rangle$ is a pseudo normed space, $x\in X$ the if we have $\{x_n\}_{n\in\{m,\ldots,\infty\}}$ defined by $x_n=x$ then $\lim_{n\rightarrow\infty} x_n=x$

Proof. Let $\varepsilon>0$ then if $n\geqslant m$ we have $\|x_n-x_n\|=\|x-x\|=0<\varepsilon$ □

The next theorem proves that the definition of convergence in a normed space is equivalent to convergence in the associated metric space (and thus by 12.306 equivalent with convergence in a topological space).

Theorem 12.314. *If $\langle X, \|\|\rangle$ is a pseudo-normed space and $\langle X, d\rangle$ the associated metric space (where $d: X\times X\rightarrow\mathbb{R}$ is defined by $d(x,y)=\|x-y\|$ (see 12.73)) then $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ is convergent to x in the normed sense iff it is convergent in the metric sense to x.*

Proof.

1. **(⇒)** Let $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ converges to x in the normed sense, then $\forall \varepsilon \in \mathbb{R}_+$ there exists a $N\in\{m,\ldots,\infty\}$ such that $\forall n \geqslant N$ we have $\|x-x_n\|<\varepsilon \Rightarrow d(x, x_n)<\varepsilon$ proving that $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ converges to x in the metric sense.

2. **(⇐)** Let $\{x_i\}_{i\in\{m,\ldots.,\infty\}}$ converges to x in the metric sense, then $\forall \varepsilon \in \mathbb{R}_+$ there exists a $N\in\{m,\ldots,\infty\}$ such that $\forall n \geqslant N$ we have $d(x,x_n)<\varepsilon \Rightarrow \|x-x_n\|<\varepsilon$ proving that $\{x_i\}_{i\in\mathbb{N}}$ converges to x in the normed sense. □

Theorem 12.315. *If $\langle X, \|\|\rangle$ is a normed space then if $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ is convergent it has a unique limit. This unique limit is noted (like in the metric case) as* $\lim_{i\to\infty} x_i$

Proof. This follows from the fact that a normed space is a metric space and in metric spaces limits are unique (see 12.307) □

Definition 12.316. *If $\langle X, \|\|\rangle$ is a pseudo-normed space then $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ is a Cauchy sequence if $\forall \varepsilon \in \mathbb{R}_+$ there exists a $N\in\{m,\ldots,\infty\}$ such that if $n,k\geqslant N$ then $\|x_n-x_k\|<\varepsilon$*

Again it is easy proved that the definition of Cauchy in a normed space is equivalent with the definition in the associated metric space

Theorem 12.317. *If $\langle X, \|\|\rangle$ is a pseudo-normed space and $\langle X, d\rangle$ the associated metric space (where $d: X\times X\rightarrow\mathbb{R}$ is defined by $d(x,y)=\|x-y\|$ (see 12.73)) then $\{x_i\}_{i\in\{m,\ldots,\infty\}}$ is Cauchy in the metric sense iff it is Cauchy in the metric sense*

Proof. We have

$$\begin{aligned}
\{x_i\}_{i\in\{m,\ldots,\infty\}}\,\text{is Cauchy (metric sense)} &\Leftrightarrow \forall\varepsilon\in\mathbb{R}_+\vDash\exists N\in\{m,\ldots,\infty\}\vdash\forall n,\\
&\quad\;\; k\geqslant N\,\text{we have}\,\|x_n-x_k\|<\varepsilon\\
&\underset{d(x,y)=\|x-y\|}{\Leftrightarrow} \forall\varepsilon\in\mathbb{R}_+\vDash\exists N\in\{m,\ldots,\infty\}\vdash\forall n,\\
&\quad\;\; k\geqslant N\,\text{we have}\,d(x_n,x_m)<\varepsilon\\
&\Leftrightarrow \{x_i\}_{i\in\{m,\ldots,\infty\}}\,\text{is Cauchy (normed sense)}
\end{aligned}$$

□

Theorem 12.318. *If $\langle X, \|\|\rangle$ is a pseudo-normed space then if $\{x_i\}_{i\in\mathbb{N}}$ is convergent then it is Cauchy.*

Proof. This follows from the fact that $\langle X, \|\|\rangle$ forms a metric space and 12.311. □

Example 12.319. If $a\in\mathbb{R}, b\in\mathbb{R}$ with $0\leqslant b<1$ then $\{a\cdot b^n\}_{n\in\mathbb{N}_0}$ converges to 0.

Proof. Two cases must be considered:

1. $\boldsymbol{(b=0)}$ then $a\cdot b^n=0$ so $\{a\cdot b^n\}_{n\in\mathbb{N}_0}$ converges to 0

2. **$(0<b)$** Let $\varepsilon \in \mathbb{R}_+$ then by 9.66 there exists an $N \in \mathbb{N}_0$ such that if $n \geqslant N$ then we have that $0 < b^n < \frac{\varepsilon}{a} \Rightarrow 0 < a \cdot b^n < \varepsilon \Rightarrow |a \cdot b^n - 0| < \varepsilon$ proving that $\{a \cdot b^n\}_{n \in \mathbb{N}_0}$ converges to 0. □

Example 12.320. Let $a \in \mathbb{R}_+$ then $\left\{\frac{1}{a+i}\right)_{i \in \mathbb{N}_0}$ converges to 0

Proof. Let $\varepsilon \in \mathbb{R}_+$ then using 9.55 there exists a $N \in \mathbb{N}$ such that $\frac{1}{\varepsilon} - a < N$ so if $N \leqslant i$ then $\frac{1}{\varepsilon} - a < i \Rightarrow \frac{1}{\varepsilon} < i + a$ so that $\frac{1}{a+i} < \varepsilon$ proving that $\left|\frac{1}{a+i} - 0\right| = \left|\frac{1}{a+i}\right| = \frac{1}{a+i} < \varepsilon$ which proves that $\lim_{i \rightarrow \infty} \frac{1}{a+i} = 0$ □

Example 12.321. Let $a \in \mathbb{R}, b \in \mathbb{R}_+$ then $\left\{\frac{a+i}{b+i}\right)_{i \in \mathbb{N}_0}$ converges to 1

Proof. First let $i \in \mathbb{N}_0$ then $\frac{a+i}{b+i} - 1 = \frac{a+i-(b+i)}{b+i} = \frac{a+b}{b+i}$. Let $\varepsilon \in \mathbb{R}_+$ then using 9.55 there exists a $N \in \mathbb{N}$ such that $\frac{|a+b|}{\varepsilon} - b < N$ then if $N \leqslant i$ we have $\frac{|a+b|}{\varepsilon} - b < i \Rightarrow \frac{|a+b|}{\varepsilon} < b + i \Rightarrow \frac{|a+b|}{b+i} < \varepsilon$ so that $\left|\frac{a+i}{b+i} - 1\right| = \left|\frac{a+b}{b+i}\right| = \frac{|a+b|}{b+i} < \varepsilon$ proving that $\lim_{i \rightarrow \infty} \frac{a+i}{b+i} = 1$

□

Theorem 12.322. *Let $\{x_i\}_{\in \{k, \ldots, \infty\}} \subseteq \mathbb{C}$ be a convergent sequence of complex numbers then $\{\bar{x_i}\}_{i \in \{k, \ldots., \infty\}}$ is convergent and $\lim_{i \rightarrow \infty} \bar{x_i} = \overline{\lim_{i \rightarrow \infty} x_i}$*

Proof. Let $\varepsilon > 0$ then there exists a $N \in \{k, \ldots, \infty\}$ such that $\forall i \geqslant N$ we have $|x_i - x| < \varepsilon$. Now $|\bar{x_i} - \bar{x}| = |\overline{x_i - x}| \underset{9.33}{=} |x_i - x| \in \varepsilon$ □

Theorem 12.323. *Let $\langle X, d_X \rangle$, $\langle Y, d_Y \rangle$ be metric spaces, $f: X \rightarrow Y$ a function and $x \in X$ then we have the following equivalence*

f is continuous at x $\Leftrightarrow$ *for every convergent sequence $\{x_n\}_{n \in \mathbb{N}}$ with limit x we have that $\{f(x_n)\}_{n \in \mathbb{N}}$ converges with limit $f(x)$*

Proof.

$\Rightarrow$. Let $\varepsilon > 0$ then by continuity of f there exists a $\delta(x) > 0$ such that if $d_X(x, y) < \lambda(x)$ then $d_Y(f(x), f(y)) < \varepsilon$. As $\lim_{n \rightarrow \infty} x_n = x$ there exists a $N \in \{k, \ldots \infty\}$ such that if $n \geqslant N$ then $d(x, x_n) < \delta(x) \Rightarrow d_Y(f(x), f(y)) < \varepsilon$ proving that $\lim_{n \rightarrow \infty} f(x_n) = f(x)$

$\Leftarrow$. Assume that f is not continuous at x then there exists a $\varepsilon_x > 0$ such that $\forall \delta > 0$ there exists a $y \in X$ with $d_X(x, y) < \delta$ and $d_Y(f(x), f(y)) > \varepsilon_x$. So $\forall n \in \mathbb{N}$ we have $\exists x_n \in X$ with $d_X(x, x_n) < \frac{1}{N}$ and $d_Y(f(x), f(y_n)) > \varepsilon_X$. Take now $\varepsilon > 0$ then by the Archimedean propety of the real numbers (see 9.55) there exists a $N_\varepsilon \in \mathbb{N}$ such that $0 < \frac{1}{N_\varepsilon} < \varepsilon$, if now $n \geqslant N_\varepsilon$ then we have $d_X(x, x_n) < \frac{1}{n} \leqslant \frac{1}{N_\varepsilon} < \varepsilon$, which proves that $\lim_{n \rightarrow \infty} x_n = x$. As we have however $\forall n \in \mathbb{N}$ that $d_Y(f(x), f(x_n)) \geqslant \varepsilon_x$ we do not have that $\lim_{n \rightarrow \infty} f(x) = f(y)$ contradicting our hypothese, so f must be continuous at x. □

Lemma 12.324. *Let $\langle X, \|\|\rangle$ be a normed space and $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ be a Cauchy sequence then there exists a $K\in\mathbb{R}_0^+$ so that $\forall i\in\{k,\ldots,\infty\}$ we have $\|x_i\|<K$*

Proof. Take $\varepsilon=1$ then because of Cauchy there exists a $N\in\{k,\ldots,\infty\}$ such that $\forall n,m\geqslant N$ we have $\|x_n-x_m\|<1$ and thus is $n\geqslant N$ we have $\|x_n\|\leqslant\|x_n-x_N\|+\|x_N\|<1+\|x_N\|$ so if we take $K=\max(\|x_1\|,\ldots,\|x_{N-1}\|,\|x_N\|+1)$ then we have $\forall i\in\{k,\ldots,\infty\}$ that $\|x_i\|<K$. □

Corollary 12.325. *Let $\langle X, \|\|\rangle$ be a normed space and $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ be a convergent sequence then there exists a $K\in\mathbb{R}_0^+$ so that $\forall i\in\{k,\ldots,\infty\}$ we have $\|x_i\|<K$.*

Proof. If $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ convergence then by 12.318 $\{x_n\}_{n\in\{,\ldots,\infty\}}$ is Cauchy. Using the previous lemma [see 12.324] we have then that there exists a $K\in\mathbb{R}_0^+$ so that $\forall i\in\{k,\ldots,\infty\}$ we have $\|x_i\|<K$. □

12.11.2 Complete spaces

Definition 12.326. *A pseudo metric space $\langle X, d\rangle$ is complete if every Cauchy sequence is convergent. A normed space $\langle X, \|\|\rangle$ is complete if every Cauchy sequence is convergent (this is the same as being Cauchy in the metric sense). A complete normed space is called a* ***Banach*** *space.*

Definition 12.327. *Let X be a topological space, $\langle Y, \|\|\rangle$ a normed space then a function $f: X\rightarrow Y$ is* ***bounded*** *iff $f(X)$ is bounded (see 12.65). In other words if $\exists M\in\mathbb{R}$ such that $\forall x,y\in X$ we have $\|f(x)-f(y)\|\leqslant M$. Let $\mathcal{B}(X,Y)\subseteq Y^X$ be defined by $\mathcal{B}(X,Y)=\{f\in Y^X|f \text{ is bounded}\}$ the set of bounded functions from X to Y.*

Theorem 12.328. *Let X be a non empty topological space, $\langle Y, \|\|\rangle$ a normed space then a function $f: X\rightarrow Y$ is* ***bounded*** *iff $\exists M\in\mathbb{R}$ such that $\forall x\in X$ we have $\|f(x)\|\leqslant M$*

Proof.

($\Rightarrow$) As f is bounded there exists a $M\in\mathbb{R}$ such that $\forall, x, y\in X$ we have $\|f(x)-f(y)\|\leqslant M$. As $X\neq\emptyset$ there exists a $x_0\in X$ let now $x\in X$ then $\|f(x)\|=\|f(x)-f(x_0)+f(x_0)\|\leqslant\|f(x_0)\|+\|f(x)-f(x_0)\|\leqslant\|f(x_0)\|+M$

($\Leftarrow$) If $\forall x\in X$ we have $\|f(x)\|\leqslant M$ then if $x,y\in X$ we have $\|f(x)-f(y)\|\leqslant\|f(x)\|+\|f(y)\|\leqslant 2M$ □

Theorem 12.329. *Let X be a set, $\langle Y, \|\|\rangle$ a normed space over a field $\mathbb{K}$ then $\mathcal{B}(X,Y)$ is a vector space (using the canonical sum and scalar product of functions defined in 10.113)*

Proof. As Y is a vector space we have by 10.113 that Y^X forms a vector space. Now if $\alpha,\beta\in\mathbb{K}$ and $f,g\in\mathcal{B}(X,Y)$ then there exists $M_f,M_g\in\mathbb{R}$ such that $\forall x,y\in X$ we have $\|f(x)-f(y)\|\leqslant M_f$, $\|g(x)-g(y)\|\leqslant M_g$. We have then that $\|\alpha\cdot f(x)+\beta\cdot g(x)-\alpha\cdot f(y)-\beta\cdot g(y)\|\leqslant\|\alpha\cdot(f(x)-f(y))+\beta\cdot(g(x)-g(y))\|\leqslant|\alpha|\cdot\|f(x)-f(y)\|+|\beta|\cdot\|g(x)-g(y)\|\leqslant|\alpha|\cdot M_f+|\beta|\cdot M_g$ proving that $\alpha\cdot f+\beta\cdot g\in\mathcal{B}(X,Y)$. Using 10.109 we have then that $\mathcal{B}(X,Y)$ forms a vector space. □

Definition 12.330. *Let X be a non empty set, $\langle Y, \|\|\rangle$ a normed space over a field $\mathbb{K}$ $f, g \in \mathcal{B}(X, Y)$ then $\|f\|_s$ is defined by* $\sup(\{\|f(x)\| | x \in X\})$

Theorem 12.331. *Let X be a non empty topological space, $\langle Y, \|\|\rangle$ a normed space over a field $\mathbb{K}$ then $\mathcal{B}(X, Y)$ then $\|\|_s: \mathcal{B}(X, Y) \to \mathbb{R}_+$ is well defined and forms a norm. So $\langle \mathcal{B}(X, Y), \|\|_s\rangle$ is a normed space.*

Proof. As $X \neq \emptyset$ there exists a $x \in X$ then if $y \in X$ we have that $\|f(y)\| \leqslant \|f(x) - f(y)\| + \|f(x)\| \leqslant M_f + \|f(x)\|$ proving that $\forall t \in \{\|f(y)\| | y \in X\}$ we have that $t \leqslant M_f + \|f(X)\|$ so that $M_f + \|f(x)\|$ forms a upper bound of $\{\|f(y)\| | y \in X\}$ and thus by 9.43 we have that $\sup(\{\|f(y)\| | y \in X\})$ exists. Now to proof that $\|\|_s$ is a norm

1. $\forall f \in \mathcal{B}(X, Y) \vDash \|f\|_s \geqslant 0$ is trivial by the definition.
2. If $f \in \mathcal{B}(X, Y)$ and $\alpha \in \mathbb{K}$ then $\{\|\alpha f(x)\| | x \in X\} = \{|\alpha| \cdot \|f(x)\| | x \in X\} = |\alpha| \cdot \{\|f(x)\| | x \in X\} \Rightarrow \|f\|_s = \sup(|\alpha| \cdot \{\|f(x)\| | x \in X\}) \underset{9.45}{=} |\alpha| \cdot \sup(\{\|f(x)\| | x \in X\}) = |\alpha| \cdot \|f\|_s$
3. If $f, g \in \mathcal{B}(X, Y)$ then

$$\begin{aligned}
\|f+g\|_s &= \sup(\{\|(f+g)(x)\| | x \in X\}) \\
&= \sup(\{\|f(x)+g(x)\| | x \in X\}) \\
&\leqslant \sup(\{\|f(x)\| + \|g(x)\| | x \in X\}) \text{ (see 2.172)} \\
&\leqslant \sup(\{\|f(x)\| | x \in X\} + \{\|g(x)\| | x \in X\}) \text{ (see 2.171)} \\
&\underset{9.46}{=} \sup(\{\|f(x)\| | x \in X\}) + \sup(\{\|g(x)\| | x \in X\}) \\
&= \|f\|_s + \|g\|_s
\end{aligned}$$

□

Theorem 12.332. *Let X be a non empty set, $\langle Y, \|\|\rangle$ a Banach space over a field $\mathbb{K}$ then $\langle \mathcal{B}(X, Y), \|\|_s\rangle$ is a Banach space.*

Proof. Let $\{f_n\}_{n \in \{k, \ldots, \infty\}}$ be a Cauchy sequence in $\mathcal{B}(X, Y)$ then given $\varepsilon > 0$ there exists a $N \in \mathbb{N}_0$ such that $\forall n, m$ we have $\sup(\{\|f(x) - g(x)\| | x \in X\}) = \|f_n - f_m\|_s < \varepsilon$ so that $\forall x \in X$ we have $\|f(x) - f(y)\| \leqslant \sup(\{\|f(x) - g(x)\| | x \in X\}) < \varepsilon$ or $\forall x \in \{f_n(x)\}_{n \in \{k, \ldots, \infty\}}$ is a Cauchy sequence so as Y is a Banach space there exists a $f(x) \in Y$ such that $\lim_{n \to \infty} f_n(x) = f(x)$.

So we can define a function $f: X \to Y$ defined by $x \to f(x)$. We prove now that f is bounded. So as $1 > 0$ there exists a $N(x) \in \mathbb{N}_0$ such that if $n \geqslant N(x)$ then $\|f_n(x) - f(x)\| < 1 \Rightarrow \|f(x)\| \leqslant \|f_n(x)\| + \|f_n(x) - f(x)\| < \|f_n(x)\| + 1$ of $\forall n \geqslant N(x)$ we have $\|f(x)\| < 1 + \|f_n(x)\|$. Now as $\{f_n\}_{n \in \{k, \ldots, \infty\}}$ is Cauchy we have that $\exists N \in \mathbb{N}_0$ such that if $n, m \geqslant \mathbb{N}_0$ then $\|f_n - f_m\|_s < 1$ so if $m \geqslant N$ then $\|f_m - f_N\|_s < 1 \Rightarrow \|f_m\|_s \leqslant \|f_N\|_s + \|f_m - f_N\|_s < \|f_N\|_s + 1$ or if $m \geqslant N$ and $x \in X$ then $\|f_m(x)\| \leqslant \sup(\{\|f_m(x)\| | x \in X\}) = \|f_m\|_s < \|f_N\|_s + 1$. Now given $x \in X$ if $l \geqslant \max(N(x), N)$ then $\|f(x)\| < 1 + \|f_l(x)\| < 1 + (\|f_N\|_s + 1) = 2 + \|f_N\|_s$ proving by 12.328 that f is a bounded function and thus that $f \in \mathcal{B}(X, Y)$.

Next we prove that $\lim_{n\rightarrow\infty} f_n = f$, so let $\varepsilon > 0$ then if $x \in X$ there exists a $N(x)$ such that if $n \geqslant N(x)$ then $\|f_n(x) - f(x)\| < \frac{\varepsilon}{2}$. As $\{f_n\}_{n\in\{k,\ldots,\infty\}}$ there exist a N such that $\|f_n - f_N\| < \frac{\varepsilon}{2}$ if $n \geqslant N$ so that $\forall x \in X$ we have $\|f_n(x) - f_N(x)\| \leqslant \sup(\{\|f_n(x) - f_N(x)\| | x \in X\}) = \|f_n - f_N\| < \frac{\varepsilon}{2}$. Take then $x \in X$ and $m \geqslant N$ and $n \geqslant \max(N(x), N)$ then we have that $\|f_m(x) - f(x)\| \leqslant \|f_m(x) - f_n(x) + f_n(x) - f(x)\| \leqslant \|f_m(x) - f_n(x)\| + \|f_n(x) - f(x)\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ proving that $\forall x \in X$ we have if $m \geqslant N$ that $\|f_m(x) - f(x)\| < \varepsilon$ proving that ε is a upper bound of $\{\|f_m(x) - f(x)\| | x \in X\}$ or $\|f_m - f\|_s = \sup(\{\|f_m(x) - f(x)\| | x \in X\}) < \varepsilon$ proving that $\lim_{n\rightarrow\infty} f_n = f$ and thus that $\mathcal{B}(X, Y)$ is Banach. □

Definition 12.333. *Let X be a set, $\langle Y, \|\|\rangle$ a metric space be topological spaces and $\{f_n\}_{n\in\{k,\ldots,\infty\}}$ a sequence of function from $X \rightarrow Y$ then $\{f_n\}_{n\in\{k,\ldots,\infty\}}$ **converges uniformely** to a function $f: X \rightarrow Y$ if $\forall \varepsilon \in \mathbb{R} \vdash \varepsilon > 0$ there exists a $N \in \mathbb{N}_0$ such that $\forall n \geqslant N$ and $\forall x \in X$ we have $\|f_n(x) - f(x)\| < \varepsilon$*

As a example of uniform convergence we use the set of bounded functions

Theorem 12.334. *Let X be a non empty set, $\langle Y, \|\|\rangle$ a normed space then if $\{f_n\}_{n\in\{k,\ldots,\infty\}}$ is a sequence of bounded functions then convergence in $\mathcal{B}(X, Y)$ is equivalent with uniform convergence.*

Proof.

1. **(⇒)** Given $\varepsilon > 0$ there exists a $N \in \mathbb{N}_0$ such that if $n \geqslant N$ then $\sup(\{\|f_n(x) - f(x)\| | x \in X\}) = \|f_n - f\|_s < \varepsilon \underset{\text{sup is } a \text{ upper bound}}{\Rightarrow} \forall x \in X$ we have $\|f_n(x) - f(x)\| < \varepsilon$ proving uniform convergence.

2. **(⇐)** Given $\varepsilon > 0$ there exists a $N \in \mathbb{N}_0$ such that if $n \geqslant N$ we have $\forall x \in X$ that $\|f(x) - f(x)\| < \varepsilon$ proving that ε is a upperbound of $\{\|f_n(x) - f(x)\| | x \in X\}$ so that $\|f_n - f\|_s = \sup(\{\|f_n(x) - f(x)\| | x \in X\}) < \varepsilon$ proving that $\lim_{n\rightarrow\infty} f_n = f$. □

Theorem 12.335. *Let $\langle X, \|\|_X\rangle$ be a normed space, $\langle Y, \|\|_Y\rangle$ a normed space then if $\{f_n\}_{n\in\{k,\ldots,\infty\}}$ is a sequence of continuous functions from X to Y uniformly converging to a function $f: X \rightarrow Y$ then f is continuous.*

Proof. Take $x \in X$ and $\varepsilon > 0$, from uniform convergence there exists a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $\forall x \in X$ that $\|f_n(x) - f(x)\|_Y < \frac{\varepsilon}{3}$. Now as $\forall i \in \{k, \ldots, \infty\}$ we have that f_N is continuous $\delta_N > 0$ such that if $\|x - y\|_X < \delta_N$ then $\|f_N(x) - f_N(y)\|_Y < \frac{\varepsilon}{3}$. Let $\|x - y\|_X < \delta_N$ then $\|f(x) - f(y)\|_Y \leqslant \|f(x) - f_N(x) + f_N(x) - f_N(y) + f_N(y) - f(y)\|_Y \leqslant \|f(x) - f_N(x)\|_Y + \|f_N(x) - f_N(y)\|_Y + \|f_N(y) - f_N(y)\|_Y < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$ proving that f is continuous. □

Theorem 12.336. *Let $\langle X, d\rangle$ is complete and $A \subseteq X$ is closed then $\langle A, d_A\rangle$ is complete.*

Proof. Let $\{x_n\}_{n\in\{k,\ldots\infty\}}$ be a Cauchy sequence in A then trivially $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ is a Cauchy sequence in X so there exists a $x\in X$ such that $\lim_{n\rightarrow\infty}x_n=x$. So if U is a open set with $x\in U$ then there exists a $\varepsilon>0$ such that $x\in B_d(x,\varepsilon)\subseteq U$ and as $\lim_{n\rightarrow\infty}x_n=x$ there exists a $N\in\{k,\ldots\infty\}$ so that $d(x,x_n)<\varepsilon$ is $n\geqslant N$ proving that $x_N\in B_d(x,\varepsilon)\subseteq U\Rightarrow U\bigcap A\neq\emptyset$ proving that $x\in A$ (see 12.21) and thus that $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ is convergent in S. □

Theorem 12.337. *Let $\langle X,\|\|\rangle$ be a normed space and let $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ be a convergent sequence with $\lim_{n\rightarrow\infty}x_n$ then if there exists a $m\in\{k,\ldots,\infty\}$ such that $\forall i\in\{m,\ldots,\infty\}$ we have $\|x_i\|\leqslant A\in\mathbb{R}$ (or $A\leqslant\|x_i\|$) then we have that $\|x\|\leqslant A$ (or $A\leqslant\|x\|$)*

Proof. Suppos that there exists a $m\in\{k,\ldots,\infty\}$ such that

1. **($\forall i\in\{m,\ldots\infty\}$ we have $\|x_i\|\leqslant A$)** We proceed by contradiction so assume that $A<\|x\|$ then there exists a $\varepsilon=\|x\|-A$ with $\varepsilon<0$ so there exists a $N\in\{k,\ldots\infty\}$ such that $\|x-x_{\max(N,m)}\|<\varepsilon$ and thus $\|x\|\leqslant\|x-x_{\max(N,m)}\|+\|x_{\max(N,m)}\|<\varepsilon+\|x_{\max(N,m)}\|\leqslant\varepsilon+A=\|x\|-A+A=\|x\|$ giving the contradiction $\|x\|<\|x\|$.
2. **($\forall i\in\{m,\ldots,\infty\}$ we have $A\leqslant\|x_i\|$)** We proceed by contradiction so assume that $\|x\|<A$ then $\varepsilon=A-\|x\|>0$ and there exists a $N\in\{k,\ldots,\infty\}$ such that $\|x_{\max(N,m)N}-x\|<\varepsilon$ then $A\leqslant\|x_{\max(N,m)}\|\leqslant\|x-x_{\max(N,m)}\|+\|x\|<\varepsilon+\|x\|=A-\|x\|+\|x\|=A$ leading to the contradiction $A<A$ so we must have $A\leqslant\|x\|$. □

Theorem 12.338. *Let $\langle X,\|\|\rangle$ be a normed space on $\mathbb{K}$, $\{x_i\}_{i\in\{k,\ldots,n\}}$ a convergent sequence in X, $\{s_i\}_{i\in\{k,\ldots,\infty\}}$ a convergent sequence in $\mathbb{R}$ with $\forall i\in\{k,\ldots,\infty\}$ $\|x_i\|<s_i$ (so $0\leqslant s_N$) then $\left\|\lim_{i\rightarrow\infty}x_i\right\|<\lim_{i\rightarrow\infty}s_i$*

Proof. Let $x=\lim_{i\rightarrow\infty}x_i$, $s=\lim_{i\rightarrow\infty}s_i$ (using the previous theorem we must then have that $0\leqslant s$ as $0\leqslant\|x_n\|\leqslant s$) and assume that $s<\|x\|$ so that $\varepsilon=\|x\|-s>0$ and find $N_1,N_2\in\{k,\ldots\infty\}$ such that $\|x_n-x\|<\frac{\varepsilon}{2}$ if $n\geqslant N_1$ and $|s_m-s|<\frac{\varepsilon}{2}$ then if $N=\max(N_1,N_2)$ we have $\|x\|\leqslant\|x-x_N\|+\|x_N\|<\frac{\varepsilon}{2}+\|x_N\|\leqslant\frac{\varepsilon}{2}+s_N\leqslant\frac{e}{2}+|s_N|\leqslant\frac{\varepsilon}{2}+|s_N-s|+|s|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}+|s|=\varepsilon+|s|\underset{0\leqslant s\Rightarrow s=|s|}{=}\varepsilon+s=\|x\|-s+s=\|x\|$ leading to the contradiction $\|x\|<\|x\|\Rightarrow\|x\|\leqslant s$. □

Theorem 12.339. *Let $\langle X,\|\|\rangle$ be a normed space and let $\{x_n\}_{n\in\{k,\ldots\infty\}}$, $\{y_n\}_{n\in\{k,\ldots,\infty\}}$ be convergent sequences with $\forall n\in\{k,\ldots,\infty\}$ we have $\|x_n\|\leqslant\|y_n\|$ then $\left\|\lim_{n\rightarrow\infty}x_n\right\|\leqslant\left\|\lim_{n\rightarrow\infty}y_n\right\|$*

Proof. Let $x=\lim_{n\rightarrow\infty}x_n$, $y=\lim_{n\rightarrow\infty}y_n$ and assume that $\|y\|<\|x\|$ so that $0<\|x\|-\|y\|=\varepsilon$ and find $N_1,N_2\in\{k,\ldots\infty\}$ such that $\|x_n-x\|<\frac{\varepsilon}{2}$,$\|y_m-y\|<\frac{\varepsilon}{2}$ if $n\geqslant N_1$, $m\geqslant N_2$ then if $N=\max(N_1,N_2)$ we have $\|x\|\leqslant\|x_N-x\|+\|x_N\|\leqslant\|x_N-x\|+\|y_N\|\leqslant\|x_N-x\|+\|y_N-y+y\|\leqslant\|x_N-x\|+\|y_N-y\|+\|y\|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}+\|y\|=\varepsilon+\|y\|=\|x\|-\|y\|+\|y\|=\|x\|\Rightarrow\|x\|<\|x\|$ a contradiction so we must have that $\|x\|\leqslant\|y\|$ □

Theorem 12.340. *Let $\langle X, \|\|\rangle$ be a normed space, $a \in X$ and $\{x_n\}_{n \in \{k,\ldots,\infty\}}$ a sequence in X then $\{x_n)n \in \mathbb{N}$ convergence if and only if $\{x_n + a\}_{n \in \mathbb{N}}$ converges. Further if we have convergence then $\lim_{n \to \infty}(x_n + a) = \left(\lim_{n \to \infty} x_n\right) + a$*

Proof.

1. Assume that $\{x_n\}_{n \in \{k,\ldots,\infty\}}$ is convergent with $\lim_{n \to \infty} x_n = x$ then if $\varepsilon > 0$ there exists a $N_\varepsilon \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N_\varepsilon$ we have $\|x_n - x\| < \varepsilon$. Take then $\varepsilon > 0$ then for $n \geqslant N_\varepsilon$ we have $\|(x_n + a) - (x + a)\| = \|x_n - x\| < \varepsilon$ which proves that $\{x_n + a\}_{n \in \mathbb{N}}$ converge with $\lim_{n \to \infty}(x_n + a) = x + a$
2. Assume that $\{x_n + a\}_{n \in \{k,\ldots,\infty\}}$ is convergent with $\lim_{n \to \infty}(x_n + a) = y$ then by (1) $\{(x_n + a) + (-a)\}_{n \in \mathbb{N}} = \{x_n\}_{n \in \mathbb{N}}$ converges with $\lim_{n \to \infty} x_n = \lim_{n \to \infty}((x_n + a) - a) \underset{(1)}{=} \lim_{n \to \infty}(x_n + a) - a = y - a$ so that $\lim_{n \to \infty}(x_n + a) = \left(\lim_{n \to \infty} x_n\right) + a$. □

Theorem 12.341. *Let $\langle X, \|\|\rangle$ be a normed space and given two sequences $\{x_n\}_{n \in \{k,\ldots,\infty\}}$, $\{y_n\}_{n \in \{k,\ldots,\infty\}}$ be convergent sequences with $\lim_{n \to \infty} x_n = x$ and $\lim_{n \to \infty} y_n \to y$ then $\{x_n + y_n\}_{n \in \{k,\ldots,\infty\}}$ is convergent with $\lim_{n \to \infty}(x_n + y_n) = x + y$. Second $\forall \alpha \in \mathbb{K}$ ($\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$ the field of scalars in X) we have $\lim_{n \to \infty}(\alpha \cdot x_n) = \alpha \cdot x$. Third if $z \in X$ then $\lim_{n \to \infty}(z + x_n) = z + \lim_{n \to \infty} x_n$.*

Proof.

1. Given $\varepsilon > 0$ find a $N_1, N_2 \in \{k, \ldots, \infty\}$ such that $n \in \{k, \ldots, \infty\}$ with $n \geqslant N_1$, N_2 we have $\|x_n - x\| < \frac{\varepsilon}{2}$, $\|y_n - y\| \leqslant \frac{\varepsilon}{2}$ so that if $n \in \{k, \ldots, \infty\}$ with $k \geqslant N = \max(N_1, N_2)$ we have $\|(x_n + y_n) - (x + y)\| \leqslant \|x_n - x\| + \|y_n - y\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$
2. We have two cases to consider
   a. $(\boldsymbol{\alpha = 0})$ then $\alpha \cdot x_i = 0$ and it is trivial to prove that $\lim_{n \to \infty}(\alpha \cdot x_n) = 0 = \alpha \cdot \lim_{n \to \infty} x_n$
   b. $(\boldsymbol{\alpha \neq 0})$ then $|\alpha| \neq 0$ and given $\varepsilon > 0$ find a $N \in \{k, \ldots, \infty\}$ so that $\forall \in \{k, \ldots, \infty\}$ with $n \geqslant N$ we have $\|x_n - x\| < \frac{\varepsilon}{|\alpha|}$ so that $\|\alpha \cdot x_n - \alpha \cdot x\| = \|\alpha \cdot (x_n - x)\| = |\alpha| \cdot \|x_n - x\| < \frac{\varepsilon}{|\alpha|} \cdot |\alpha| = \varepsilon$
3. Given $\varepsilon > 0$ then there exists a $N \in \{k, \ldots, \infty\}$ such that if $n \geqslant \mathrm{N}_0$, $n \in \{k, \ldots, \infty\}$ that $\|x_n - x\| < \varepsilon$ then $\|(z + x_n) - (z + x)\| = \|x_n - x\| < \varepsilon$. □

Theorem 12.342. *Let $\{x_i\}_{i \in \{k,\ldots,\infty\}} \subseteq [0, \infty[$ be a convergent sequence of positive numbers then $0 \leqslant \lim_{i \to \infty} x_i$*

Proof. Assume that $x = \lim_{i \to \infty} x_i < 0$ then $0 < -x = \varepsilon$ then there exists a $N \in \{k, \ldots, \infty\}$ such that $|x_N - x| < \frac{\varepsilon}{2}$. As $0 \leqslant x_N, x < 0 \Rightarrow 0 < -x$ we have that $0 < x_N - x$ so that $x_N - x = |x_N - x| < \frac{\varepsilon}{2} \Rightarrow x_N < \frac{\varepsilon}{2} + x = \frac{\varepsilon}{2} - \varepsilon = -\frac{\varepsilon}{2} < 0$ contradicting $0 \leqslant x_N$ so we must have that $0 \leqslant \lim_{i \to \infty} x_i$ □

Corollary 12.343. *Let $\{x_i\}_{i\in\{k,\ldots,\infty\}}, \{y_i\}_{i\in\{k,\ldots n\}}\subseteq\mathbb{R}$ be convergent sequences with $\forall i\in\{k,\ldots,\infty\}$ $x_i\leqslant y_i$ then $\lim_{i\rightarrow\infty} x_i\leqslant\lim_{i\rightarrow\infty} y_i$*

Proof. Using the assumption we have that $\{y_i-x_i\}_{i\in\{k,\ldots,\infty\}}\subseteq[0,\infty[$ hence using the previous theorem we have that $\lim_{i\rightarrow\infty} y_i-\lim_{i\rightarrow\infty} x_i=\lim_{i\rightarrow\infty}(y_i-x_i)\geqslant 0$ hence $\lim_{i\rightarrow\infty} y_i\geqslant\lim_{i\rightarrow\infty} x_i$. □

Theorem 12.344. *Let X be a set with two equivalent compatible norms $\|\|_1$, $\|\|_2$ then*

1. *$\{x_i\}_{i\in\{k,\ldots,\infty\}}$ is convergent to $x\in X$ in $\|\|_1$ if and only if it is convergent to x in $\|\|_2$*
2. *$\{x_i\}_{i\in\{k,\ldots,\infty\}}$ is Cauchy in $\|\|_1$ if and only if it is Cauchy in $\|\|_2$*

Proof. First note that we only have to prove $\Rightarrow$ (as $\Leftarrow$ follows from applying the theorem in the opposite direction. Next as $\|\|_1$ and $\|\|_2$ are equivalent there exists $M_1>0, M_2>0$ such that $\forall x\in X$ we have $M_1\cdot\|x\|_1\leqslant\|x\|_2\leqslant M_2\cdot\|x\|_1$. We have then

1. **(Convergence)** Let $\varepsilon>0$ then by convergence in $\|\|_1$ there exists a $N\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N$ we have $\|x-x_n\|<\frac{\varepsilon}{M_2}$ so that $\|x-x_n\|_2\leqslant M_2\cdot\|x-x_n\|_1<M_2\cdot\frac{\varepsilon}{M_2}=\varepsilon$
2. **(Cauchy)** Let $\varepsilon>0$ then by the Cauchy property in $\|\|_1$ there exists a $N\in\{k,\ldots,\infty\}$ such that $\forall n,m\geqslant N$ we have $\|x_n-x_m\|<\frac{\varepsilon}{M_2}$ and thus $\|x_n-x_m\|_2\leqslant M_2\cdot\|x_n-x_m\|_1<M_2\cdot\frac{e}{M_2}=\varepsilon$ □

Theorem 12.345. *Let $\langle X,d\rangle$ be a metric space, $\{x_i\}_{i\in\{k,\ldots\infty\}}$ a sequence and x a limit point of $\{x_i|i\in\{,\ldots,\infty\}\}$ then $\forall N\in\{k,\ldots,\infty\}$, $\varepsilon>0$ there exists a $n\in\{k,\ldots,\infty\}$ with $n>N$ such that $d(x,x_n)<\varepsilon$*

Proof. Given $\varepsilon>0$ take then $A_\varepsilon=\{i\in\{k,\ldots,\infty\}|0<d(x,x_i)<\varepsilon\}\subseteq\{k,\ldots,\infty\}$. We prove now that A_ε must be finite by contradiction. So assume that A_ε is finite. Then as x is a limit point of $\{x_i|i\in\{k,\ldots,\infty\}\}$ there exists a $x_i\in\{x_i|i\in\{k,\ldots,\infty\}|\}$ with $x_i\neq x$ and $x\in B_d(x,\varepsilon)\Rightarrow 0<d(x,x_i)<\varepsilon\Rightarrow i\in A_\varepsilon\Rightarrow A_\varepsilon\neq\emptyset$. As we assumed that A_ε is finite, we have that $B_\varepsilon=\{d(x,x_i)|i\in A_\varepsilon\}$ is finite (as $f: A_\varepsilon\rightarrow B_\varepsilon$, $i\rightarrow f(i)=d(x,x_i)$ is a surjection) and we have by 5.50 the existence of a $\delta=\min(B_\varepsilon)=\min(d(x,x_i)|i\in A_\varepsilon|)$. We must have that $\delta>0$ (as $\forall i\in A_\varepsilon$ we have $d(x,x_i)>0$ by definition of A_ε). Again by the fact that x is the limit point of $\{x_i|i\in\{k,\ldots,\infty\}|\}$ that there exists a $x_j\in\{x_i|i\in\{k,\ldots,\infty\}\}$ such that $x_j\neq x$ and $x_j\in B_d(x,\delta)\Rightarrow 0<d(x,x_j)<\delta\Rightarrow j\in A_\varepsilon\Rightarrow d(x,x_j)\in B_\varepsilon=\{d(x,x_i)|i\in A_\varepsilon\}\Rightarrow\delta\leqslant d(x,x_j)<\delta\Rightarrow\delta<\delta$ a contradiction. So we must have that A_ε is infinite.

If the theorem is not true then $\exists N_0\in\{k,\ldots,\infty\}$ such that $\forall n>N_0$ that $d(x,x_n)\geqslant\varepsilon\Rightarrow n\notin A_\varepsilon$ and thus $A_\varepsilon\subseteq\{k,\ldots,N_0\}$ so A_ε would be finite contradicting the fact that A_ε is infinite. □

Theorem 12.346. *Let $\langle X,d\rangle$ be a metrics space then every Cauchy sequence has a maximum of one limit point.*

Proof. Assume that x, y are different limit points of the Cauchy sequence $\{x_i\}_{i \in \{k, \ldots, \infty\}}$. Then as $x \neq y$ we have that $\varepsilon = d(x, y) > 0$. Now by the Cauchy condition we can find a $N \in \{k, \ldots, \infty\}$ such that $\forall i, j \geqslant N$ we have $d(x_i, x_j) < \frac{\varepsilon}{3}$. By 12.345 there exists $n, m \in \{k, \ldots, \infty\}$ such that $n, m > N$ and $d(x, x_n) < \frac{\varepsilon}{2}$, $d(x, x_m) < \frac{\varepsilon}{2} \Rightarrow \varepsilon = d(x, y) \leqslant d(x, x_n) + d(x_n, x_m) + d(y, x_m) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \Rightarrow \varepsilon < \varepsilon$ a contradiction so we must have that $x = y$. □

12.11.3 Examples of complete spaces

Theorem 12.347. *Every compact metric space $\langle X, d\rangle$ is complete.*

Proof. Let $\{x_i\}_{i \in \{k, \ldots, \infty\}}$ be a arbitrary Cauchy sequence and let $S = \{x_i | i \in \{k, \ldots \infty\}\}$ then we must consider the following cases:

1. **(S is finite)** In this case there exists a $N \in \{k, \ldots, \infty\}$ such that $\forall i, j \geqslant N$ we have $x_i = x_j$
2. [[We prove this by contradiction, so assume that $\forall N \in \{k, \ldots, \infty\}$ there exists $i, j \geqslant N$ such that $x_i \neq x_j$ so that $A = \{d(x, y) | x, y \in S \wedge x \neq y\}$ is not empty (take $N = k$ then there exists a $i, j \geqslant k$ with $x_i \neq x_j$) and finite (as by 5.44 $S \times S$ is finite and $f: S \times S \rightarrow A$ defined by $(x, y) \rightarrow d(x, y)$ is a surjection). By 5.50 $\varepsilon = \min(A)$ exists and is >0 (for if $d \in A \Rightarrow d = d(x, y) > 0$ as $x \neq y$). By the Cauchy option there exists then a $N_1 \in \{k, \ldots, \infty\}$ such that if i', $j' \geqslant N_1$ then $d(x_i, x_j) < \varepsilon$ and by the hypothesis there exists a $i, j \geqslant N_1$ such that $x_i \neq x_j \Rightarrow d(x_i, x_j) \in A \Rightarrow \varepsilon \leqslant d(x, y) < \varepsilon \Rightarrow \varepsilon < \varepsilon$ a contradiction]]. Given a $\varepsilon > 0$ take $n \geqslant N$ then $x_n = x_N \Rightarrow d(x_n, x_N) = 0 < \varepsilon$ proving that $\{x_i\}_{i \in \{k, \ldots, \infty\}}$ converges to x_N.
3. **(S is infinite)** Then using 12.253 there exists limit points for S which because of the previous theorem must be unique (see 12.346), let's call this limit point x. Take now $\varepsilon > 0$ then by the Cauchy property there exist a $N \in \{k, \ldots, \infty\}$ such that $\forall i, j \geqslant N$ we have $d(x_i, x_j) < \frac{\varepsilon}{2}$, by 12.345 there exists a $n \in \{k, \ldots, \infty\}$ such that $n \geqslant N$ and $d(x, x_n) < \frac{\varepsilon}{2}$ so that if $m \geqslant N$ then $d(x, x_m) \leqslant d(x, x_n) + d(x_n, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ proving convergence to x. □

Theorem 12.348. *The normed space $\langle \mathbb{R}, ||\rangle$ of the real numbers together with the default metric is complete.*

Proof. Let $\{x_i\}_{i \in \{k, \ldots, \infty\}}$ be a Cauchy sequence in $\mathbb{R}$ so that there exists a $N \in \{k, \ldots, \infty\}$ such that $\forall i, j \geqslant N$ that $|x_i - x_j| < 1$. Take then $A = \max(|x_1|, \ldots, |x_{N-1}|, |x_N| + 1)$ and $[-A, A]$. If $i \geqslant N$ we have $|x_i| \leqslant |x_i - x_N| + |x_N| < 1 + |x_N| \leqslant A \Rightarrow x_i \in [-A, A]$ and if $i < N \Rightarrow |x_i| \leqslant A$ so that $\forall i \in \{k, \ldots, \infty\}$ we have $x_i \in [-A, A]$. By 12.257 $[-A, A]$ is compact in $\mathbb{R}$ (and thus compact in the subspace topology)) and by the previous theorem (see 12.347) $[-A, A]$ is complete in the subspace topology and thus as $\{x_n\}_{n \in \{k, \ldots, \infty\}}$ is trivially Cauchy in the subspace topology there exists a $x \in [-A, A]$ such that $\exists M \in \{k, \ldots, \infty\}$ such that $\forall m \geqslant M$ we have $|x - x_m| = |x - x_m|_{|A} < \varepsilon$ proving convergence to x. □

Corollary 12.349. *The normed space $\langle \mathbb{C}, ||_c\rangle$ of the complex numbers together with the complex norm is complete. Here $||_c$ is the canoncial norm in the complex space in contrast with the absolute value $||$.*

Proof. Let $\{z_i\}_{i\in\{k,\ldots,\infty\}}$ be a Cauchy sequence, where $\forall i\in\{k,\ldots,\infty\}$ we have $z_i=x_i+i\cdot y_i$. Then $\forall n,m\in\{k,\ldots,\infty\}$ we have

$$\begin{array}{rcl} |x_m-x_n| & = & \sqrt{(x_m-x_n)^2} \\ & \leqslant_{9.71} & \sqrt{(x_m-x_n)^2+(y_m-y_n)^2} \\ & = & |z_m-z_n|_c \end{array}$$

and like wise

$$\begin{array}{rcl} |y_m-y_n| & = & \sqrt{(y_m-y_n)^2} \\ & \leqslant_{9.71} & \sqrt{(x_m-x_n)^2+(y_m-y_n)^2} \\ & = & |z_m-z_n|_c \end{array}$$

proving that $\forall n,m\in\{k,\ldots,n\}\vDash|x_m-x_n|,|y_m-y_n|\leqslant|z_m-z_n|_c$. Given a $\varepsilon>0$ there exists a $N\in\{k,\ldots,\infty\}$ so that $\forall n,m\geqslant N$ we have that $|z_n-z_m|_c<\varepsilon\Rightarrow|x_m-x_n|$, $|y_m-y_n|<\varepsilon$. So we have that $\{x_i\}_{i\in\{k,\ldots,\infty\}},\{y_i\}_{i\in\{k,\ldots,\infty\}}$ is Cauchy and as $\langle\mathbb{R},||\rangle$ is complete $x=\lim_{i\rightarrow\infty}x_i, y=\lim_{i\rightarrow\infty}y_i$ exists. So given a $\varepsilon>0$ there exists $N_1,N_2\in\{k,\ldots,\infty\circ\}$ such that if $k\geqslant N_1, l\geqslant N_2$ that $|x_n-x|<\frac{\varepsilon}{2},|y_n-y|<\frac{\varepsilon}{2}$. Then $|z_m-z_n|_c=|(x_m-x_n)+i\cdot(y_m-y_n)|\leqslant|x_m-x_n|_c+|y_m-y_n|_c=|x_m-x_n|+|y_m-y_n|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$. This proves that $\lim_{i\rightarrow\infty}z_i$ exists and that $\langle\mathbb{C},||_c\rangle$ is complete. □

Theorem 12.350. *Let $\langle\mathbb{R}^n,\|\|\rangle$ be the normed space based on the maximum (product) norm (see 12.79) $\|x\|=\max(\{|x_i||i\in\{1,\ldots,n\}\})$ then $\langle\mathbb{R}^n,\|\|\rangle$ is a Banach space. As all norms on $\mathbb{R}^n$ are equivalent (see 12.293) we have that $\langle\mathbb{R}^n,\|\|\|\|\rangle$ is a Banach space for every norm $\|\|\|\|$ on $\mathbb{R}^n$.*

Proof. If $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ is a Cauchy sequence then by the previous lemma there exist a $K\in\mathbb{R}_0^+$ such that $\forall i\in\{k,\ldots,\infty\}$ we have $\|x_i\|<K\Rightarrow\|x_i\|\leqslant K$ or if we take $\overline{B_{d_{\|\|}}}(0,K)=\{x\in\mathbb{R}^n|\|x\|\leqslant K\}$ then $\forall i\in\{k,\ldots,\infty\}$ we have $x_i\in B$. $\overline{B_{d_{\|\|}}}(0,K)$ is closed (see 12.56) and trivially bounded [if $x,y\in\overline{B_{d_{\|\|}}}(0,K)$ then $d_{\|\|}(x,y)=\|x-y\|\leqslant\|x\|+\|y\|\leqslant K+K=2\cdot K$] so by the Heine-Borel theorem (12.290) we have that it is compact and thus by 12.347 is complete. So the sequence $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ in $\overline{B_{d_{\|\|}}}(0,K)$ being Cauchy in $\overline{B_{d_{\|\|}}}(0,K)$ has a limit x which is trivially also the limit of $\{x_i\}_{i\in\{k,\ldots,n\}}$. □

Theorem 12.351. *Every finite dimensional normed vector space is a Banach space*

Proof. Let $\langle X,\|\|\rangle$ be a n-dimensional normed vector space then using 12.169 there exists a norm $\|\|$ on $\mathbb{R}^n$ and a isometry (also isomorphism) φ between X and $\mathbb{R}^n$. If then $\{x_i\}_{i\in\{k,\ldots\infty\}}$ is a Cauchy sequence in X we have for $\varepsilon>0$ the existence of a $N\in\{k,\ldots,\infty\}$ such that if $n,m\geqslant N$ that $\|x_n-x_m\|<\varepsilon$. Take now the sequence $\{\varphi(x_i)\}_{i\in\{k,\ldots,\infty\}}$ then if $n,m\geqslant N$ we have that $\|\varphi(x_n)-\varphi(x_m)\|=\|\varphi(x_n-x_m)\|\underset{\varphi\text{ is a isometry}}{=}\|x_n-x_m\|<\varepsilon$ proving that $\{\varphi(x_i)\}_{i\in\{k,\ldots,\infty\}}$ is Cauchy, using the previous theorem we have that $\langle\mathbb{R}^n,\|\|\rangle$ is a Banach space and so $\{\varphi(x_i)\}_{i\in\{k,\ldots,\infty\}}$ has a limit $y\in\mathbb{R}^n$ or in other words if $\varepsilon>0$ then there exists a $N\in\{k,\ldots,\infty\}$ such that if $n\geqslant N$ then $\|\varphi(x_i)-y\|<\varepsilon$. If we then take $x=\varphi^{-1}(y)$ then if $n\geqslant N$ we have that $\|x_n-x\|=\|\varphi(x_n-x)\|=\|\varphi(x_n)-\varphi(x)\|=\|\varphi(x_n)-\varphi(\varphi^{-1}(y))\|=\|\varphi(x_n)-y\|<\varepsilon$. This finally proves that $\langle X,\|\|\rangle$ is a Banach space. □

Theorem 12.352. *Let* $\{\langle X_i, \|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ *be a family of normed spaces and let* $\langle Y, \|\|\rangle$ *be a Banach space then* $L(X_1, \ldots, X_n; Y)$ *is a Banach space using the operator norm.*

Proof. Let $\{L_i\}_{i\in\{k,\ldots,\infty\}}$ be a Cauchy sequence in $L(X_1, \ldots, X_n; Y)$ and take $(x_1, \ldots, x_n) \in \prod_{i\in\{k,\ldots,\infty\}} X_i$ then we have for $p, q \in \{k, \ldots, \infty\}$

$$\|L_p(x_1, \ldots, x_n) - L_q(x_1, \ldots, x_n)\| = \|(L_p - L_q)(x_1, \ldots, x_n)\| \leqslant \|L_p - L_q\| \cdot \prod_{i\in\{1,\ldots,n\}} \|x_i\|_i \tag{12.28}$$

Take now the family $\{L_p(x_1, \ldots, x_n)\}_{p\in\{k,\ldots,\infty\}}$ and take $\varepsilon > 0$ then we have two possible cases:

1. $\left(\prod_{i\in\{1,\ldots,n\}} \|x_i\|_i = 0\right)$ then by 12.28 we have that $\|L_p(x_1, \ldots, x_n) - L_q(x_1, \ldots, x_n)\| = 0 < \varepsilon$ so if $p, q \geqslant k \in \{k, \ldots, \infty\}$ then $\|L_p(x_1, \ldots., x_n) - L_q(x_1, \ldots, x_n)\| < \varepsilon$
2. $\left(\prod_{i\in\{1,\ldots,n\}} \|x_i\|_i \neq 0\right)$ take then $\varepsilon' = \frac{\varepsilon}{\prod_{i\in\{1,\ldots,n\}} \|x_i\|_i}$ and as $\{L_p\}_{p\in\{k,\ldots,\infty\}}$ is Cauchy there exists a $N \in \{k, \ldots, \infty\}$ such that if $p, q \geqslant N$ then $\|L_p - L_q\| < \varepsilon' \underset{12.28}{\Rightarrow} \|L_p(x_1, \ldots, x_n) - L_q(x_1, \ldots, x_n)\| \leqslant \|L_p - L_q\| \cdot \prod_{i\in\{1,\ldots,n\}} \|x_i\|_i < \varepsilon' \cdot \prod_{i\in\{1,\ldots,n\}} \|x_i\|_i = \varepsilon$

from (1) and (2) we conclude then that

$$\{L_p(x_1, \ldots, x_n)\}_{p\in\{k,\ldots,\infty\}} \text{ is } a \text{ Cauchy sequence in } Y \tag{12.29}$$

As $\langle Y, \|\|\rangle$ is a Banach space it follows from 12.29 that $\{L_p(x_1, \ldots, x_n)\}_{p\in\{k,\ldots,\infty\}}$ has a limit (depending of course on $(x_1, \ldots, x_n)$) let's note this limit as $L(x_1, \ldots, x_n)$ so we have

$$L(x_1, \ldots, x_n) = \lim_{q\to\infty} L_p(x_1, \ldots, x_n) \tag{12.30}$$

Define now the function $L: \prod_{i\in\{1,\ldots,n\}} X_i \to Y$ by $(x_1, \ldots, x_n) \to L(x_1, \ldots, x_n)$. We prove now that $L \in L(X_1, \ldots, X_n; Y)$:

1. **(multilinearity)** Given $\alpha, \beta \in \mathbb{K}$, $(x_1, \ldots, x_n) \in \prod_{i\in\{1,\ldots,n\}} X_i$, $i \in \{1, \ldots, n\}$ and $x, y \in X_i$ then we have that $L(x_1, \ldots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_n) = \lim_{p\to\infty} L_p(x_1, \ldots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \ldots, x_n) \underset{L_p \text{ is multilinear}}{=} \lim_{p\to\infty} (\alpha \cdot L_p(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) + \beta \cdot L_p(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)) \underset{12.341}{=} \alpha \cdot \lim_{p\to\infty} L_p(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) + \beta \cdot \lim_{p\to\infty} L_p(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) = \alpha \cdot L(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) + \beta \cdot L(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)$

2. **(continuity)** Take the family $\{\|L_p\|\}_{p\in\{k,\ldots,\infty\}}$ in $\mathbb{R}$ then if $\varepsilon>0$ we have from the fact that $\{L_p\}_{p\in\{k,\ldots\infty\}}$ is Cauchy there exists a $N\in\{k,\ldots\infty\}$ such that if $p,q\geqslant N$ then $\|L_p-L_q\|<\varepsilon\Rightarrow|\|L_p\|-\|L_q\||\leqslant\|L_p-L_q\|<\varepsilon$ proving that $\{\|L_p\|\}_{p\in\{k,\ldots,\infty\}}$ is Cauchy in $\langle\mathbb{R},||\rangle$ and as $\langle\mathbb{R},||\rangle$ is complete (see 12.348) there exists a $A\in\mathbb{R}$ with $A=\lim_{p\to\infty}\|L_p\|$. Using 12.339 and $\|L_p(x_1,\ldots,x_n)\|\leqslant\|L_p\|\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i$ we have that $\|L(x_1,\ldots x_n)\|=\lim_{p\to\infty}\|L_p(x_1,\ldots,x_n)\|\leqslant\lim_{p\to\infty}\left(\|L_p\|\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|\right)=\left(\lim_{p\to\infty}\|L_p\|\right)\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i=A\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i$ proving that L is indeed continuous and that

$$\|L\|\leqslant A=\lim_{p\to\infty}\|L_p\| \tag{12.31}$$

So the only thing left to prove is that $\{L_q\}_{q\in\{k,\ldots\infty\}}$ converges to L. Let $\varepsilon>0$ then for every $(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}X_i$ with $\forall i\in\{1,\ldots,n\}$ $\|x_i\|=1$ we have by the fact that $\{L_q\}_{q\in\{1,\ldots,\infty\}}$ is Cauchy there exists a $N_1\in\{k,\ldots,\infty\}$ such that if p, $q\geqslant N_1$ then $\|L_p-L_q\|<\frac{\varepsilon}{2}$ so that $\|L_p(x_1,\ldots,x_n)-L_q(x_1,\ldots,x_n)\|=\|(L_p-L_q)(x_1,\ldots,x_n)\|\leqslant\|L_p-L_q\|\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i=\|L_p-L_q\|<\frac{\varepsilon}{4}$. Further using 12.30 there exists a $N_2(x_1,\ldots x_n)\in\{k,\ldots,\infty\}$ such that if $r\geqslant N_2(x_1,\ldots,x_n)$ then $\|L(x_1,\ldots,x_n)-L_r(x_1,\ldots,x_n)\|<\frac{\varepsilon}{4}$. So that if $p\geqslant N_1$ then we have $\|(L-L_p)(x_1,\ldots,x_n)\|=\|L(x_1,\ldots,x_n)-L_p(x_1,\ldots,x_n)\|\leqslant\|L(x_1,\ldots,x_n)-L_s(x_1,\ldots,x_n)\|+\|L_s(x_1,\ldots,x_n)-L_p(x_1,\ldots,x_n)\|<\frac{\varepsilon}{4}+\frac{\varepsilon}{4}=\frac{\varepsilon}{2}$ if we choose $s\geqslant\max(N_1,N_2(x_1,\ldots,x_n))$ proving that if $p\geqslant N_1$ then $\|(L-L_p)(x_1,\ldots,x_n)\|<\frac{\varepsilon}{2}$ proving by 12.191 that $\|L-L_p\|\leqslant\frac{\varepsilon}{2}<\varepsilon$ and thus that $\lim_{p\to\infty}L_p=L$. □

Corollary 12.353. *Let $\langle X,\|\|_X\rangle$ be a normed space, $\langle Y,\|\|_Y\rangle$ a Banach space, $n\in\mathbb{N}$ then $\langle L^n(X;Y),\|\|_{L^n(X;Y)}\rangle$ is a Banach space (using the operator norms)*

Proof. Let $S=\{n\in\mathbb{N}|$If $\langle X,\|\|_X\rangle$ is a normed space, $\langle Y,\|\|_Y\rangle$ a Banach space then $L^n(A;Y)$ is a Banach space$\}$ then:

1. If $n=1$ then $L^1(X;Y)=L(X,Y)=L(X;Y)$ is a Banach space by the above theorem, so that $1\in S$.
2. If $n\in S$ then $L^n(X;Y)$ is a Banach space (if $\langle X,\|\|_X\rangle$ is a normed space, $\langle Y,\|\|_Y\rangle$ a Banach space) so that by (1) $L^{n+1}(X;Y)=L(X;L^n(X;Y))$ is a Banach space proving that $n+1\in S$

Induction proves then the theorem. □

For the real numbers we have for increasing [or decreasing] sequences the following result

Lemma 12.354. *Let $\{x_i\}_{i\in\{k,\ldots,\infty\}}\subseteq\mathbb{R}$ then we have*

1. *If $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ is increasing [$\forall i\in\{k,\ldots,\infty\}$ we have $x_i\leqslant x_{i+1}$] then*
   a. *If $\sup(\{x_i|i\in\{k,\ldots,\infty\}\})$ exists then $\lim_{i\to\infty}x_i$ exists and $\lim_{i\to\infty}x_i=\sup(\{x_i|i\in\{k,\ldots,\infty\}\})$*

b. *If* $\lim_{i\rightarrow\infty} x_i$ *exists then* $\sup(\{x_i|i\in\{k,\ldots,\infty\}\})$ *exists and* $\lim_{i\rightarrow\infty} x_i = \sup(\{x_i|i\in\{k,\ldots,\infty\}\})$

2. *If* $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ *is decreasing [*$\forall i\in\{k,\ldots,\infty\}$ *we have* $x_{i+1}\leqslant x_i$*] then*

   a. *If* $\inf(\{x_i|i\in\{k,\ldots,\infty\}\})$ *exists then* $\lim_{i\rightarrow\infty} x_i$ *exists and* $\lim_{i\rightarrow\infty} x_i = \inf(\{x_i|i\in\{k,\ldots,\infty\}\})$

   b. *If* $\lim_{i\rightarrow\infty} x_i$ *exists then* $\inf(\{x_i|i\in\{k,\ldots,\infty\}\})$ *exists and* $\lim_{i\rightarrow\infty} x_i = \inf(\{x_i|i\in\{k,\ldots,\infty\}\})$

Proof.

1. Assume that $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ is increasing then

   a. Let $x=\sup(\{x_i|i\in\{k,\ldots,\infty\}\})$ and take $\varepsilon>0$ then $x-\varepsilon<x<x+\varepsilon$ and by the definition of a supremum there exists a $N\in\{k,\ldots,\infty\}$ such that $x-\varepsilon<x_N\leqslant x<x+\varepsilon$. If now $n\geqslant N$ we have (see 12.301) that $x_N\leqslant x_n\leqslant x\Rightarrow x-\varepsilon<x_n\leqslant x<x+\varepsilon\Rightarrow|x-x_n|<\varepsilon$ proving that $\lim_{i\rightarrow\infty} x_i=x=\sup(\{x_i|i\in\{k,\ldots,\infty\}\})$

   b. As $\lim_{i\rightarrow\infty} x_i=x$ we have for $1>0$ that there exists a $N\in\{k,\ldots,\infty\}$ such that $\forall i\geqslant N$ we have $|x_i-x|<1\Rightarrow x_i-x<1\Rightarrow x_i<1+x$. Hence $\forall i\in\{k,\ldots,\infty\}$ we have $x_i\leqslant\max(\{x_i|i\in\{k,\ldots,N\}\}\bigcup\{1+x\})$. Using 9.43 we have then that $\sup(\{x_i|i\in\{k,\ldots,\infty\}\})$ exists. Let $y=\sup(\{x_i|i\in\{k,\ldots,\infty\}\})$ then using (a) that we have just proved we get that $y=\lim_{i\rightarrow\infty} x_i=x$ proving using the uniqueness of a limit that $\sup(\{x_i|i\in\{k,\ldots,\infty\}\})=x$

2. Assume that $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ is decreasing then

   a. Let $x=\inf(\{x_i|i\in\{k,\ldots,\infty\}\})$ and take $\varepsilon>0$ then $x-\varepsilon<x<x+\varepsilon$ and by the definition of a infinum there exists a $N\in\{k,\ldots,\infty\}$ such that $x-\varepsilon<x\leqslant x_N<x+\varepsilon$. If now $n\geqslant N$ we have (see 12.301) that $x\leqslant x_n\leqslant x_N\Rightarrow x-\varepsilon<x\leqslant x_n<x+\varepsilon\Rightarrow|x-x_n|<\varepsilon$ proving that $\lim_{i\rightarrow\infty} x_i=x=\inf(\{x_i|i\in\{k,\ldots,\infty\}\})$

   b. As $\lim_{i\rightarrow\infty} x_i=x$ we have for $1>0$ that there exists a $N\in\{k,\ldots,\infty\}$ such that $\forall i\geqslant N$ we have $|x_i-x|<1\Rightarrow x-x_i<1\Rightarrow x-1<x_i$. Hence $\forall i\in\{k,\ldots,\infty\}$ we have $\min(\{x_i|i\in\{k,\ldots,N\}\}\bigcup\{x-1\})\leqslant x_i$. Using 9.43 we have then that $\inf(\{x_i|i\in\{k,\ldots,\infty\}\})$ exists. Let $y=\inf(\{x_i|i\in\{k,\ldots,\infty\}\})$ then using (a) that we have just proved we get that $y=\lim_{i\rightarrow\infty} x_i=x$ proving using the uniqueness of a limit that $\inf(\{x_i|i\in\{k,\ldots,\infty\}\})=x$ □

Example 12.355. Let we have that

1. $\lim_{n\rightarrow\infty}\frac{1}{n}=0$

2. $\inf\left(\left\{\frac{1}{n}|n\in\mathbb{N}\right\}\right)=0$

Proof.

1. This is proved in 12.320.
2. As $\forall n\in\mathbb{N}$ we have $\frac{1}{n+1}\leqslant\frac{1}{n}$ $\left\{\frac{1}{n}\right\}_{n\in\mathbb{N}}$ is a decreasing seuence, hence by the previous lemma we have $\inf\left(\left\{\frac{1}{n}|n\in\mathbb{N}\right\}\right)=\lim_{n\rightarrow\infty}\frac{1}{n}=0$. □

12.11.4 Limit infinum and limit supremum in $\mathbb{R}$

For this subsection we use the normed space $\langle\mathbb{R},||\rangle$ based on the absolute value in $\mathbb{R}$

Definition 12.356. *Let $\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq\mathbb{R}$ then we say*

1. $\limsup_{n\rightarrow\infty} x_i$ *exists iff* $\forall n\in\{k,\ldots,n\}$ $\sup(\{x_i|i\in\{n,\ldots,\infty\}\})$ *exists and* $\lim_{n\rightarrow\infty}\sup(\{x_i|i\in\{n,\ldots,\infty\}\})$ *exists. In that case* $\limsup_{n\rightarrow\infty} x_i=\lim_{n\rightarrow\infty}\sup(\{x_i|i\in\{n,\ldots,\infty\}\})$.
2. $\liminf_{n\rightarrow\infty} x_i$ *exists iff* $\forall n\in\{k,\ldots,n\}$ $\inf(\{x_i|i\in\{n,\ldots,\infty\}\})$ *exists and* $\lim_{n\rightarrow\infty}\inf(\{x_i|i\in\{n,\ldots,\infty\}\})$ *exists. In that case* $\liminf_{n\rightarrow\infty} x_i=\lim_{n\rightarrow\infty}\inf(\{x_i|i\in\{n,\ldots,\infty\}\})$.

Proposition 12.357. *Let $\{x_i\}_{i\in\{k,\ldots,\infty\}}\subseteq\mathbb{R}$ then*

1. $\limsup_{n\rightarrow\infty} x_i$ *exists if and only if* $\forall n\in\{k,\ldots,n\}$ $\sup(\{x_i|i\in\{n,\ldots,\infty\}\})$ *exists and* $\inf(\{\sup(\{x_i|i\in\{n,\ldots,\infty\}\})|n\in\{k,\ldots,\infty\}\})$ *exists. Further if* $\limsup_{n\rightarrow\infty} x_i$ *exists then* $\limsup_{n\rightarrow\infty} x_i=\inf(\{\sup(\{x_i|i\in\{n,\ldots,\infty\}\})|n\in\{k,\ldots,\infty\}\})$.
2. $\liminf_{n\rightarrow\infty} x_i$ *exists if and only if* $\forall n\in\{k,\ldots,n\}$ $\inf(\{x_i|i\in\{n,\ldots,\infty\}\})$ *exists and* $\sup(\{\inf(\{x_i|i\in\{n,\ldots,\infty\}\})|n\in\{k,\ldots,\infty\}\})$ *exists. Further if* $\liminf_{n\rightarrow\infty} x_i$ *exists then* $\liminf_{n\rightarrow\infty} x_i=\sup(\{\inf(\{x_i|i\in\{n,\ldots,\infty\}\})|n\in\{k,\ldots,\infty\}\})$

Proof.

1. If $\forall n\in\{k,\ldots,\infty\}$ $\sup(\{x_i|i\in\{n,\ldots,\infty\}\})$ exists then $\forall n\in\{k,\ldots,\infty\}$ we have $\sup(\{x_i|i\in\{n,\ldots,\infty\}\})\leqslant\sup(\{x_i|i\in\{n+1,\ldots,\infty\}\})$. Applying then 12.354 proves (1)
2. If $\forall n\in\{k,\ldots,\infty\}$ $\inf(\{x_i|i\in\{n,\ldots,\infty\}\})$ exists then $\forall n\in\{k,\ldots,\infty\}$ we have $\inf(\{x_i|i\in\{n+1,\ldots,\infty\}\})\leqslant\inf(\{x_i|i\in\{n+1,\ldots,\infty\}\})$. Applying then 12.354 proves (1) □

Example 12.358. Let $x\in\mathbb{R}$ then if we define $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ by $x_i=x$ then $\liminf_{n\rightarrow\infty} x_i,\limsup_{n\rightarrow\infty} x_i$ exists and $x=\liminf_{n\rightarrow\infty} x_i=\limsup_{n\rightarrow\infty} x_i$

Proof. Let $n \in \{k, \ldots, \infty\}$ then $\sup(\{x_i | i \in \{n, \ldots, \infty\}\}) = \sup(\{x\}) = x$ and $\inf(\{x_i | i \in \{1, \ldots, n\}\}) = \inf(\{x\}) = x$ and thus $\inf(\{\sup(\{x_i | i \in \{n, \ldots, \infty\}\}) | n \in \{k, \ldots, \infty\}\}) = \inf(\{x\}) = x$ and $\sup(\{\inf(\{x_i | i \in \{n, \ldots, \infty\}\}) | n \in \{k, \ldots, \infty\}\}) = \sup(\{x\}) = x$ which by the previous proposition [see 12.357] proves that $\liminf_{n\rightarrow\infty} x_i$,$\limsup_{n\rightarrow\infty} x_i$ exists and $x = \liminf_{n\rightarrow\infty} x_i = \limsup_{n\rightarrow\infty} x_i$ □

Proposition 12.359. *Let $\{x_i\}_{i\in\{k,\ldots,\infty\}} \subseteq \mathbb{R}$ be such that $\liminf_{n\rightarrow\infty} x_i$ and $\limsup_{n\rightarrow\infty} x_i$ exists then $\liminf_{n\rightarrow\infty} x_n \leqslant \limsup_{n\rightarrow\infty} x_n$*

Proof. Fix $n \in \{k, \ldots, \infty\}$ then $\forall m \in \{k, \ldots, \infty\}$ we have either

$m \in \{k, \ldots, n-1\}$. then $m \leqslant n$ so that $\{x_i | i \in \{n, \ldots, \infty\}\} \subseteq \{x_i | i \in \{m, \ldots, \infty\}\}$ hence $\inf(\{x_i | i \in \{m, \ldots, \infty\}\}) \leqslant \inf(\{x_i | i \in \{n, \ldots, \infty\}\}) \leqslant x_n \leqslant \sup(\{x_i | i \in \{n, \ldots, \infty\}\})$

$m \in \{n, \ldots, \infty\}$. then $\inf(\{x_i | i \in \{m, \ldots, \infty\}\}) \leqslant x_n \leqslant \sup(\{x_i | i \in \{n, \ldots, \infty\}\})$

so we have

$$\forall m \in \{k, \ldots, \infty\} \text{ we have } \inf(\{x_i | i \in \{m, \ldots, \infty\}\}) \leqslant x_n \leqslant \sup(\{x_i | i \in \{n, \ldots, \infty\}\})$$

and thus

$$\sup(\{\inf(\{x_i | i \in \{m, \ldots, \infty\}\}) | m \in \{k, \ldots, \infty\}\}) \leqslant \sup(\{x_i | i \in \{n, \ldots, \infty\}\})$$

giving that

$$\sup(\{\inf(\{x_i | i \in \{m, \ldots, \infty\}\}) | m \in \{k, \ldots, \infty\}\}) \leqslant \inf(\{\sup(\{x_i | i \in \{n, \ldots, \infty\}\}) | n \in \{k, \ldots, \infty\}\})$$

using then 12.357 we have

$$\liminf_{n\rightarrow\infty} x_n \leqslant \limsup_{n\rightarrow\infty} x_n$$

proving the theorem. □

Proposition 12.360. *Let $\{x_i\}_{i\in\{k,\ldots,\infty\}} \subseteq \mathbb{R}$ then $\lim_{n\rightarrow\infty} x_i$ exists if and only if $\limsup_{n\rightarrow\infty} x_i$ exists and $\liminf_{n\rightarrow\infty} x_i$ exists and $\liminf_{n\rightarrow\infty} x_i = \limsup_{n\rightarrow\infty} x_i$. If $\lim_{n\rightarrow\infty} x_i$ exists then $\lim_{n\rightarrow\infty} x_i = \liminf_{n\rightarrow\infty} x_i = \limsup_{n\rightarrow\infty} x_i$.*

Proof.

1. Assume that $x = \lim_{n\rightarrow\infty} x_i$ exists then by 12.325 there exists a $K \in \mathbb{R}_0^+$ so that $\forall i \in \{k, \ldots, \infty\}$ we have $|x_i| < K$ so that $-x_i, x_i < K$ giving

$$\forall i \in \{k, \ldots, \infty\} \text{ we have } -K < x_i < K \tag{12.32}$$

Let $n \in \{k, \ldots, \infty\}$ then $\{x_i | i \in \{n, \ldots, \infty\}\}$ is bounded above by K and below by $-K$. Using the conditional completeness of $\mathbb{R}$ [see 9.43] we have

$$\forall n \in \{k, \ldots, \infty\} \quad \sup(\{x_i | i \in \{n, \ldots, \infty\}\}) \text{ and } \inf(\{x_i | i \in \{n, \ldots, \infty\}\}) \text{ exists} \tag{12.33}$$

and using the definition of the supremum and infinum we have

$$-K \leqslant \sup(\{x_i|i \in \{n, ..., \infty\}\}) \text{ and } \inf(\{x_i|i \in \{n, ..., \infty\}\}) \leqslant K \tag{12.34}$$

Using the above we have that $\{\sup(\{x_i|i \in \{n, ..., \infty\}\})|n \in \{k, ..., \infty\}\}$ is bounded below by $-K$ and $\{\inf(\{x_i|i \in \{n, ..., \infty\}\})|n \in \{k, ..., \infty\}\}$ is bounded above by K hence we have that

$$\inf(\{\sup(\{x_i|i \in \{n, ..., \infty\}\})|n \in \{k, ..., \infty\}\}) \text{ and } \sup(\{\inf(\{x_i|i \in \{n, ..., \infty\}\})|n \in \{k, ..., \infty\}\}) \text{ exists} \tag{12.35}$$

Using then 12.357 we have that

$$\limsup_{n \to \infty} x_n \text{ and } \liminf_{n \to \infty} x_n \text{ exists} \tag{12.36}$$

Assume now that $x < \sup(\{\inf(\{x_i|i \in \{n, ..., \infty\}\})|n \in \{k, ..., \infty\}\})$ then there exists a $l \in \{k, ..., \infty\}$ such that $x < \inf(\{x_i|i \in \{l, ..., \infty\}\})$. Take $\varepsilon = \inf(\{x_i|i \in \{l, ..., \infty\}\}) - x$ then $\varepsilon > 0$. Hence there exists a $N \in \{k, ..., \infty\}$ such that $\forall i \in \{N, ..., \infty\}$ we have $|x - x_i| < \varepsilon$. Take $M = \max(l, N)$ then we have $x - \varepsilon < x_M < x + \varepsilon = x + \inf(\{x_i|i \in \{l, ..., \infty\}\}) - x = \inf(\{x_i|i \in \{l, ..., \infty\}\}) \leqslant x_M$ leading to the contradiction $x_M < x_M$. Hence we must have

$$\sup(\{\inf(\{x_i|i \in \{n, ..., \infty\}\})|n \in \{k, ..., \infty\}\}) < x = \lim_{n \to \infty} x_n \tag{12.37}$$

Assume now that $\inf(\{\sup(\{x_i|i \in \{n, ..., \infty\}\})|n \in \{k, ..., \infty\}\}) < x$ then $\exists l \in \{k, ..., \infty\}$ such that $\sup(\{x_i|i \in \{l, ..., \infty\}\}) < x$. Take $\varepsilon = x - \sup(\{x_i|i \in \{l, ..., \infty\}\})$ then $0 < \varepsilon$ and there exists a $N \in \{k, ..., \infty\}$ such that $\forall i \in \{N, ..., \infty\}$ we have $|x - x_i| < \varepsilon$. Take $M = \max(N, l)$ then we have $x + \varepsilon > x_M > x - \varepsilon = x - x + \sup(\{x_i|i \in \{l, ..., \infty\}\}) = \sup(\{x_i|i \in \{l, ..., \infty\}\}) \geqslant x_M$ leading to the contradiction $x_M > x_M$. Hence we must have

$$\lim_{n \to \infty} x_n = x \leqslant \sup(\{\inf(\{x_i|i \in \{n, ..., \infty\}\})|n \in \{k, ..., \infty\}\}) \tag{12.38}$$

Finally we have $\lim_{n \to \infty} x_n \leqslant_{12.38} \sup(\{\inf(\{x_i|i \in \{n, ..., \infty\}\})|n \in \{k, ..., \infty\}\}) \leqslant_{12.359} \inf(\{\sup(\{x_i|i \in \{n, ..., \infty\}\})|n \in \{k, ..., \infty\}\}) \leqslant \lim_{n \to \infty} x_i$ which proves that

$$\lim_{n \to \infty} x_i = \liminf_{n \to \infty} x_i = \limsup_{n \to \infty} x_i$$

2. Assume that $\limsup_{n \to \infty} x_i$ exists and $\liminf_{n \to \infty} x_i$ exists and $\liminf_{n \to \infty} x_i = \limsup_{n \to \infty} x_i$. Then using 12.357 we have that $x = \liminf_{n \to \infty} x_i = \limsup_{n \to \infty} x_i = \sup(\{\inf(\{x_i|i \in \{n, ..., \infty\}\})|n \in \{k, ..., \infty\}\}) \leqslant_{12.359} \inf(\{\sup(\{x_i|i \in \{n, ..., \infty\}\})|n \in \{k, ..., \infty\}\})$. Take $\varepsilon > 0$ then we have as $x - \varepsilon < x < x + \varepsilon$ that there exists N_1, $N_2 \in \{k, ..., \infty\}$ such that $x - \varepsilon < \inf(\{x_i|i \in \{N_1, ..., \infty\}\})$ and $\sup(\{(x_i|i \in \{N_2, ..., \infty\})\}) < x + \varepsilon$. Take $N = \max(N_1, N_2)$ then $\forall i \in \{N, ..., \infty\}$ we have $x - \varepsilon < \inf(\{x_i|i \in \{N_1, ..., \infty\}\}) \leqslant x_i \leqslant \sup(\{(x_i|i \in \{N_2, ..., \infty\})\}) < x + \varepsilon$ proving that $|x - x_i| < \varepsilon$. So $\lim_{n \to \infty} x_i$ exists and is equal to $\liminf_{n \to \infty} x_i$ and $\limsup_{n \to \infty} x_i$.

□

Lemma 12.361. *Let* $\{x_i\}_{i\in\{k,\ldots.,\infty\}}, \{y_i\}_{i\in\{k,\ldots,\infty\}} \subseteq \mathbb{R}$ *the we have*

1. *If* $\sup(\{x_i|i\in\{k,\ldots,\infty\}\}), \sup(\{y_i|i\in\{k,\ldots,\infty\}\})$ *exists then* $\sup(\{x_i+y_i|i\in\{k,\ldots,\infty\}\})$ *exists and* $\sup(\{x_i+y_i|i\in\{k,\ldots,\infty\}\}) \leqslant \sup(\{x_i|i\in\{k,\ldots,\infty\}\}) + \sup(\{y_i|i\in\{k,\ldots,\infty\}\})$
2. *If* $\inf(\{x_i|i\in\{k,\ldots,\infty\}\}), \inf(\{y_i|i\in\{k,\ldots,\infty\}\})$ *exists then* $\inf(\{x_i+y_i|i\in\{k,\ldots,\infty\}\})$ *exists and* $\inf(\{x_i+y_i|i\in\{k,\ldots,\infty\}\}) \geqslant \inf(\{x_i|i\in\{k,\ldots,\infty\}\}) + \inf(\{y_i|i\in\{k,\ldots,\infty\}\})$

Proof.

- Take $i\in\{k,\ldots,\infty\}$ then $x_i+y_i \leqslant \sup(\{x_i|i\in\{k,\ldots,\infty\}\}) + \sup(\{y_i|i\in\{k,\ldots,\infty\}\})$ we have that $\{x_i+y_i|i\in\{k,\ldots,\infty\}\}$ is bouned above by $\sup(\{x_i|i\in\{k,\ldots,\infty\}\}) + \sup(\{y_i|i\in\{k,\ldots,\infty\}\})$, as $\mathbb{R}$ is conditionally complete we have that $\sup(\{x_i+y_i|i\in\{k,\ldots,\infty\}\})$ exists and

  $$\sup(\{x_i+y_i|i\in\{k,\ldots,\infty\}\}) \leqslant \sup(\{x_i|i\in\{k,\ldots,\infty\}\}) + \sup(\{y_i|i\in\{k,\ldots,\infty\}\})$$

- Take $i\in\{k,\ldots,\infty\}$ then $\inf(\{x_i|i\in\{k,\ldots,\infty\}\}) + \inf(\{y_i|i\in\{k,\ldots,\infty\}\}) \leqslant x_i+y_i$ so that s bouned below by $\inf(\{x_i|i\in\{k,\ldots,\infty\}\}) + \inf(\{y_i|i\in\{k,\ldots,\infty\}\})$, as $\mathbb{R}$ is conditionally complete we have that $\inf(\{x_i+y_i|i\in\{k,\ldots,\infty\}\})$ exists and

  $$\inf(\{x_i|i\in\{k,\ldots,\infty\}\}) + \inf(\{y_i|i\in\{k,\ldots,\infty\}\}) \leqslant \inf(\{x_i+y_i|i\in\{k,\ldots,\infty\}\})$$

□

Proposition 12.362. *Take* $\{x_i\}_{i\in\{k,\ldots.\infty\}}$ *then we have*

1. *If* $\liminf_{n\rightarrow\infty} x_i$ *exists then* $\limsup_{n\rightarrow\infty}(-x_i)$ *exists and* $\liminf_{n\rightarrow\infty} x_i = -\limsup_{n\rightarrow\infty}(-x_i)$
2. *If* $\limsup_{n\rightarrow\infty} x_i$ *exists then* $\liminf_{n\rightarrow\infty}(-x_i)$ *exists and* $\limsup_{n\rightarrow\infty} x_i = -\liminf_{n\rightarrow\infty}(-x_i)$
3. *If* $\limsup_{n\rightarrow\infty} x_i$ *exists and* $\alpha\in[0,\infty[$ *then* $\limsup_{n\rightarrow\infty}(\alpha\cdot x_i)$ *exists and* $\limsup_{n\rightarrow\infty}(\alpha\cdot x_i) = \alpha\cdot\limsup_{n\rightarrow\infty} x_i$
4. *If* $\liminf_{n\rightarrow\infty} x_i$ *exists and* $\alpha\in[0,\infty[$ *then* $\liminf_{n\rightarrow\infty}(\alpha\cdot x_i)$ *exists and* $\liminf_{n\rightarrow\infty}(\alpha\cdot x_i) = \alpha\cdot\liminf_{n\rightarrow\infty} x_i$
5. *If* $x\in\mathbb{R}$ *and* $\liminf_{n\rightarrow\infty} x_i$ *exists then for* $\liminf_{n\rightarrow\infty}(x_i+x)$ *exists and* $\liminf_{n\rightarrow\infty}(x_i+x) = \liminf_{n\rightarrow\infty} x_i + x$

6. *If* $x \in \mathbb{R}$ *and* $\limsup_{n \rightarrow \infty} x_i$ *exists then for* $\{x_i + x\}_{i \in \{k, \ldots, \infty\}}$ $\limsup_{n \rightarrow \infty} (x_i + x)$ *exists and* $\limsup_{n \rightarrow \infty} (x_i + x) = \limsup_{n \rightarrow \infty} x_i + x$

Proof.

1. As $\liminf_{n \rightarrow \infty} x_i$ exists we have

$$\forall n \in \{k, \ldots, \infty\} \inf (\{x_i | i \in \{n, \ldots, \infty\}\}) \text{ exists} \tag{12.39}$$

Using 9.44 we have then that $\sup (\{-x_i | i \in \{n, \ldots, \infty\}\})$ exists and

$$\forall n \in \mathbb{N} \in \{k, \ldots, \infty\} \sup (\{-x_i | i \in \{n, \ldots, \infty\}\}) = -\inf (\{x_i | i \in \{n, \ldots, \infty\}\}) \tag{12.40}$$

Again usingthe existance of $\liminf_{n \rightarrow \infty} x_i$ we have the existance of $\lim_{n \rightarrow \infty} \inf(\{x_i | i \in \{n, \ldots, \infty\}\})$, further we have

$$\begin{aligned}
\lim_{n \rightarrow \infty} \inf(\{x_i | i \in \{n, \ldots, \infty\}\}) & \underset{12.40}{=} \lim_{n \rightarrow \infty} (-\sup (\{-x_i | i \in \{n, \ldots, \infty\} \\
& \underset{12.341}{=} -\lim_{n \rightarrow \infty} (\sup (\{-x_i | i \in \{n, \ldots, \infty\}\})) \\
& \underset{\text{def}}{=} -\limsup_{n \rightarrow \infty} (-x_n)
\end{aligned}$$

proving (1).

2. As $\limsup_{n \rightarrow \infty} x_i$ exists we have

$$\forall n \in \{k, \ldots, \infty\} \sup (\{x_i | i \in \{n, \ldots, \infty\}\}) \text{ exists} \tag{12.41}$$

Using 9.44 we have then that $\inf (\{-x_i | i \in \{n, \ldots, \infty\}\})$ exists and

$$\forall n \in \mathbb{N} \in \{k, \ldots, \infty\} \sup (\{-x_i | i \in \{n, \ldots, \infty\}\}) = -\inf (\{x_i | i \in \{n, \ldots, \infty\}\}) \tag{12.42}$$

Again usingthe existance of $\limsup_{n \rightarrow \infty} x_i$ we have the existance of $\lim_{n \rightarrow \infty} \sup(\{x_i | i \in \{n, \ldots, \infty\}\})$, further we have

$$\begin{aligned}
\lim_{n \rightarrow \infty} \sup(\{x_i | i \in \{n, \ldots, \infty\}\}) & \underset{12.42}{=} \lim_{n \rightarrow \infty} (-\inf (\{-x_i | i \in \{n, \ldots, \infty\} \\
& \underset{12.341}{=} -\lim_{n \rightarrow \infty} (\inf (\{-x_i | i \in \{n, \ldots, \infty\}\})) \\
& \underset{\text{def}}{=} -\liminf_{n \rightarrow \infty} (-x_n)
\end{aligned}$$

proving (1).

3. As $\limsup_{n \rightarrow \infty} x_i$ exists we have

$$\forall n \in \{k, \ldots, \infty\} \sup (\{x_i | i \in \{n, \ldots, \infty\}\}) \text{ exists}$$

Using 9.45 we have that $\sup(\{\alpha \cdot x_i | i \in \{n, \ldots, \infty\}\})$ exists and is equal to $\alpha \cdot \sup(\{x_i | i \in \{n, \ldots, \infty\}\})$. So as $\lim_{n \to \infty} \sup(\{x_i | i \in \{n, \ldots, \infty\}\})$ exists we have by 12.341 that $\lim_{n \to \infty} (\alpha \cdot \sup(\{x_i | i \in \{n, \ldots, \infty\}\}))$ exists. Hence using 12.341 we have

$$\begin{array}{rcl}\alpha \cdot \limsup_{n \to \infty} x_n & = & \alpha \cdot \lim_{n \to \infty} \sup(\{\alpha \cdot x_i | i \in \{n, \ldots, \infty\}\}) \\ & \underset{12.341}{=} & \lim_{n \to \infty} (\alpha \cdot \sup(\{x_i | i \in \{n, \ldots, \infty\}\})) \\ & = & \lim_{n \to \infty} (\sup(\{\alpha \cdot x_i | i \in \{n, \ldots, \infty\}\})) \\ & = & \limsup_{n \to \infty} (\alpha \cdot x_n)\end{array}$$

4. As $\liminf_{n \to \infty} x_i$ exists we have

$$\forall n \in \{k, \ldots, \infty\} \inf(\{x_i | i \in \{n, \ldots, \infty\}\}) \text{ exists}$$

Using 9.45 we have that $\inf(\{\alpha \cdot x_i | i \in \{n, \ldots, \infty\}\})$ exists and is equal to $\alpha \cdot \inf(\{x_i | i \in \{n, \ldots, \infty\}\})$. So as $\lim_{n \to \infty} \inf(\{x_i | i \in \{n, \ldots, \infty\}\})$ exists we have by 12.341 that $\lim_{n \to \infty} (\alpha \cdot \inf(\{x_i | i \in \{n, \ldots, \infty\}\}))$ exists. Hence using 12.341 we have

$$\begin{array}{rcl}\alpha \cdot \liminf_{n \to \infty} x_n & = & \alpha \cdot \lim_{n \to \infty} \sup(\{\alpha \cdot x_i | i \in \{n, \ldots, \infty\}\}) \\ & \underset{12.341}{=} & \lim_{n \to \infty} (\alpha \cdot \sup(\{x_i | i \in \{n, \ldots, \infty\}\})) \\ & = & \lim_{n \to \infty} (\sup(\{\alpha \cdot x_i | i \in \{n, \ldots, \infty\}\})) \\ & = & \limsup_{n \to \infty} (\alpha \cdot x_n)\end{array}$$

5. As $\liminf_{n \to \infty} x_i$ exists we have that $\forall n \in \{k, \ldots, \infty\}$. Using 9.47 we have that $\sup(\{x_i + x | i \in \{n, \ldots, \infty\}\})$ exists and $\sup(\{x_i + x | i \in \{n, \ldots, \infty\}\}) = \sup(\{x_i | i \in \{n, \ldots, \infty\}\}) + x$. Hence using 12.340 we have

$$\begin{array}{rcl}\lim_{n \to \infty} (\sup(\{x_i + x | i \in \{n, \ldots, \infty\}\})) & = & \lim_{n \to \infty} (\sup(\{x_i | i \in \{n, \ldots, \infty\}\}) + x) \\ & \underset{12.340}{=} & \lim_{n \to \infty} \sup(\{x_i | i \in \{n, \ldots, \infty\}\}) + x \\ & = & \limsup_{n \to \infty} x_i + x\end{array}$$

6. As $\liminf_{n \to \infty} x_i$ exists we have that $\forall n \in \{k, \ldots, \infty\}$. Using 9.49 we have that $\inf(\{x_i + x | i \in \{n, \ldots, \infty\}\})$ exists and $\inf(\{x_i + x | i \in \{n, \ldots, \infty\}\}) = \inf(\{x_i | i \in \{n, \ldots, \infty\}\}) + x$. Hence using 12.340 we have

$$\begin{array}{rcl}\lim_{n \to \infty} (\inf(\{x_i + x | i \in \{n, \ldots, \infty\}\})) & = & \lim_{n \to \infty} (\inf(\{x_i | i \in \{n, \ldots, \infty\}\}) + x) \\ & \underset{12.340}{=} & \lim_{n \to \infty} \inf(\{x_i | i \in \{n, \ldots, \infty\}\}) + x \\ & = & \liminf_{n \to \infty} x_i + x\end{array}$$

$\square$

12.11.5 Series

The idea of series is to introduce the concept of infinite sums. Note that, in contrast with finite sums, infinte sums in general do not have the properties of commutativity and associativity. We will define infinite sums as a limit of finite sums. We use the following definitions:

Definition 12.363. (Serie) *Let $\langle X, \|\|\rangle$ be a normed space and $\{x_i\}_{i\in\{k,\ldots,\infty\}} \subseteq X$ a sequence then we note:*

1. *$\sum_{i=k}^{\infty} x_i$ is another notation for $\{x_i\}_{i\in\{k,\ldots,\infty\}}$*
2. *$\sum_{i=k}^{\infty} x_i$ is convergent if $\lim_{n\rightarrow\infty} (\sum_{i=k}^{n} x_i)$ exists.*
3. *If $\sum_{i=k}^{\infty} x_i$ converges then we note $\lim_{n\rightarrow\infty} (\sum_{i=k}^{n} x_i)$ by $\sum_{i=k}^{\infty} x_i$*

Theorem 12.364. *Let $\langle X, \|\|\rangle$ be a normed space then if $\sum_{i=k}^{\infty} x_i$ is convergent it follows that $\lim_{n\rightarrow\infty} x_n = 0$ and $\exists K \in \mathbb{R}_+$ such that $\forall i \in \{k, \ldots, \infty\}$ we have $\|x_i\| \leqslant K$*

Proof. Given $\varepsilon > 0$ then because $\{\sum_{i=k}^{n} x_i\}_{n\in\{k,\ldots,\infty\}}$ has a limit it must be Cauchy so there exists a $N \in \{k, \ldots, \infty\}$ such that if $n \geqslant N + 1$ and thus $n, n-1 \geqslant N$ then $\|x_n - 0\| = \|x_n\| = \|\sum_{i=k}^{n} x_i - \sum_{i=k}^{n-1} x_i\| < \varepsilon$ proving that $\lim_{n\rightarrow\infty} x_n = 0$. Further as $\lim_{n\rightarrow\infty} x_i$ exists we have by 12.311 that $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ is Cauchy hence using 12.324 there exists a $K \in \mathbb{R}_+$ such that $\forall i \in \{k, \ldots, \infty\}$ we have $\|x_i\| \leqslant K$. □

Theorem 12.365. *Let $\langle X, \|\|\rangle$ be a Banach space $\sum_{i=k}^{\infty} x_i$ converge if and only if $\forall \varepsilon > 0$ there exists a $N \in \{k, \ldots, \infty\}$ such that $\forall n, m \geqslant N \vdash n \leqslant m$ we have $\|\sum_{i=n}^{m} x_i\| < \varepsilon$*

Proof. Take $\{s_i\}_{i\in\{k,\ldots,\infty\}}$ defined by $s_i = \sum_{j=k}^{i} x_i$ then

⇒. If $\sum_{i=k}^{\infty} x_i$ converges we have that $\lim_{i\rightarrow\infty} s_i$ exists, take $\varepsilon > 0$, then using 12.311 there exists a $N' \in \{k, \ldots, \infty\}$ such that $\forall n, m \geqslant N'$ we have $\|s_m - s_n\| < \varepsilon$. Take $n, m \geqslant N = N' + 1$ with $n \leqslant m$ then $N' \leqslant n - 1, m$ and $\|\sum_{i=n}^{m} x_i\| \underset{10.26}{=} \|\sum_{i=k}^{m} x_i - \sum_{i=k}^{n-1} \| = \|s_m - s_n\| \leqslant \varepsilon$

⇐. Let $\varepsilon > 0$ then there exists a $N \in \{k, \ldots, \infty\}$ such that $\forall n, m \geqslant N$ with $n \leqslant m$ we have $\|\sum_{i=n}^{m} x_i\| < \varepsilon$, as $\sum_{i=n+1}^{m} x^i = \sum_{i=k}^{m} x^i - \sum_{i=k}^{n} x^i = s_m - s_n$ we conclude that $\|s_m - s_n\| = \|\sum_{n+1}^{m} x_i\| < \varepsilon$. From this it follows by definition that $\{s_i\}_{i\in\{k,\ldots,\infty\}}$ is Cauchy and as X is a Banach space and thus complete, $\lim_{i\rightarrow\infty} s_i$ exists proving that $\sum_{i=k}^{\infty} x_i$ converges. □

Theorem 12.366. *Let $\langle X, \|\|\rangle$ be a normed space, $i \in \mathbb{N}_0$, $\{x_i\}_{i\in\{k+l,\ldots,\infty\}}$ then $\sum_{i=k}^{\infty} x_{i+l}$ is convergent if and only if $\sum_{i=k+l}^{\infty} x_i$ is convergent. Further $\sum_{i=k}^{\infty} x_{i+l} = \sum_{i=k+l}^{\infty} x_i$*

Proof.

- **⇒.** Suppose $\sum_{i=k}^{\infty} x_{i+l}$ converges to x. Let $\varepsilon > 0$ then there exists a $N' \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N'$ we have $\|\sum_{i=k}^{n} x_{i+l} - x\| < \varepsilon$. Take then $N = N' + l$ then if $n \geqslant N = N' + l$ we have that $n - l \geqslant N'$. Now $\|\sum_{i=k+l}^{n} x_i - x\| = \|\sum_{i=k}^{n-l} x_{i+l} - x\| < \varepsilon$ (as $n - l \geqslant N'$) proving that $\sum_{i=k+l}^{\infty} x_i$ converges and $\sum_{i=k+l}^{\infty} x_i = \sum_{i=k}^{\infty} x_{i+l}$

- **⇐.** Suppose $\sum_{i=k+l}^{\infty}$converges to x. Let $\varepsilon > 0$ then there exists a $N' \in \{k + l, \ldots, \infty\}$ such that $\forall n \geqslant N'$ we have $\|\sum_{i=k+l}^{n} x_i - x\|$. Take $N = N' - l$ then if $n \geqslant N$ we have that $n + l \geqslant N'$. Now $\|\sum_{i=k}^{n} x_{i+l} - x\| = \|\sum_{i=k+l}^{n+l} x_i - x\| < \varepsilon$ (as $n + l \geqslant N'$) proving that $\sum_{i=k}^{\infty} x_i$ converges and $\sum_{i=k+l}^{\infty} x_i = \sum_{i=k}^{\infty} x_{i+l}$ □

Theorem 12.367. *Let $\langle X, \|\|\rangle$ be a normed space and $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ a sequence and $m \in \{k, \ldots, \infty\}$ then $\sum_{i=k}^{\infty} x_i$ is convergent if and only if $\sum_{i=m+1}^{\infty} x_i$ is a convergent series (so $\forall m \in \{k, \ldots, \infty\}$ $\sum_{i=m}^{\infty} x_i$ converges). Also if any of the two series is convergent then we have $\sum_{i=k}^{\infty} x_i = (\sum_{i=k}^{m} x_i) + (\sum_{i=m+1}^{\infty} x_i)$*

Proof.

1. **(⇒)** Given $\varepsilon > 0$ then $\exists N \in \{k, \ldots, \infty\}$ such that if $n \geqslant N$ we have $\|\sum_{i=k}^{n} x_i - \sum_{i=k}^{\infty}\| < \varepsilon$. Now take $N' = N + (m - k) + 1 \in \{m + 1, \ldots., \infty\}$ [as $N \geqslant k$] and if $n \geqslant N' \geqslant N$ [as $m \geqslant k$] then $\|\sum_{i=m+1}^{n} x_i - (\sum_{i=k}^{\infty} x_i - \sum_{i=k}^{m} x_i)\| = \|\sum_{i=m+1}^{n} x_i + \sum_{i=k}^{m} x_i - \sum_{i=k}^{\infty} x_i\| = \|\sum_{i=k}^{n} x_i - \sum_{i=k}^{\infty} x_i\| < \varepsilon$ proving that $\sum_{i=m+1}^{\infty} x_i$ exists and is equal to $\sum_{i=k}^{\infty} x_i - \sum_{i=k}^{m} x_i$

2. **(⇐)** Given $\varepsilon > 0$ then $\exists N \in \{m + 1, \ldots, \infty\}$ such that if $n \geqslant N$ we have $\|\sum_{i=m+1}^{n} x_i - \sum_{i=m+1}^{\infty} x_i\| < \varepsilon$. Then also $N \in \{k, \ldots, \infty\}$ and $\|\sum_{i=k}^{n} x_i - (\sum_{i=k}^{m} x_i + \sum_{i=m+1}^{\infty} x_i)\| = \|\sum_{i=m+1}^{n} x_i - \sum_{i=m+1}^{\infty} x_i\| < \varepsilon$ proving that $\sum_{i=k}^{\infty} x_i$ exist and is equal to $\sum_{i=k}^{m} x_i + \sum_{i=m+1}^{\infty} x_i$ □

Theorem 12.368. *Let $\langle X, \|\|\rangle$ be a normed space and $\sum_{i=k}^{\infty} x_i$ convergent serie then $\lim_{N\rightarrow\infty} \sum_{i=N}^{\infty} x_i = 0$*

Proof. Let $\varepsilon > 0$ then as $\sum_{i=k}^{\infty} x_i$ converges there exists a $N \geqslant k$ so that $\forall m - 1 \geqslant N$ we have $\|\sum_{i=k}^{\infty} x_i - \sum_{i=k}^{m-1} x_i\| < \varepsilon$ then $\|\sum_{i=m}^{\infty} x_i - 0\| = \|\sum_{i=m}^{\infty} x_i\| \underset{\text{see previous theorem 12.367}}{=} \|\sum_{i=k}^{\infty} x_i - \sum_{i=k}^{m-1} x_i\| < \varepsilon$ proving our theorem. □

Theorem 12.369. *Let $\langle X, \|\|\rangle$ be a normed space over the field $\mathbb{K}$, $\alpha, \beta \in \mathbb{K}$ and $\sum_{i=k}^{\infty} x_i$, $\sum_{i=k}^{\infty} y_i$ convergent series then $\sum_{i=k}^{\infty} (\alpha \cdot x_i + \beta \cdot y_i)$ is convergent and $\sum_{i=k}^{\infty} (\alpha \cdot x_i + \beta \cdot y_i) = \alpha \cdot \sum_{i=k}^{\infty} x_i + \beta \cdot \sum_{i=k}^{\infty} y_i$.*

Proof. Given $n \in \{k, \ldots, \infty\}$ we have that $\sum_{i=k}^{n} (\alpha \cdot x_i + \beta \cdot y_i) \underset{10.59,10.53}{=} \alpha \cdot \sum_{i=k}^{n} x_i + \beta \cdot \sum_{i=k}^{n} y_i$, using 12.341 we have then that $\lim_{n \to \infty} \sum_{i=k}^{n} (\alpha \cdot x_i + \beta \cdot y_i)$ exists and is equality to $\alpha \cdot \lim_{n \to \infty} \sum_{i=k}^{n} x_i + \beta \cdot \lim_{n \to \infty} \sum_{i=k}^{n} y_i$. □

12.11.6 Series of positive numbers

Series are in general not commutative but series of positive numbers (and later absolute convergent series) are commutative.

Lemma 12.370. *Let $\sum_{i=k}^{\infty} x_i$ be such that $\forall i \in \{k, \ldots, \infty\}$ we have $x_i \in [0, \ldots, \infty[$ then the following are equivalent*

1. $\sum_{i=k}^{\infty} x_i$ *is convergent and* $\sum_{i=k}^{\infty} x_i = x$
2. $\sup(\{\sum_{i=k}^{n} x_i | n \in \{k, \ldots, \infty\}\})$ *exists and* $x = \sup(\{\sum_{i=k}^{\infty} x_i | n \in \{k, \ldots, \infty\}\})$
3. $\sup(\{\sum_{n \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\} \wedge K \text{ is finite})\})$ *exists and* $x = \sup(\{\sum_{n \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\} \wedge K \text{ is finite})\})$
4. $\forall \varepsilon > 0$ *there exists a* $n \in \{k, \ldots, \infty\}$ *such that* $\forall p \in \mathbb{N}$ *and* $l \geqslant n$ *we have* $\sum_{i=l+1}^{l+p} x_i < \varepsilon$
5. $\forall \varepsilon > 0$ *there exists a finite* $K \subseteq \{k, \ldots, \infty\}$ *such that* $\forall H \subseteq \{k, \ldots, \infty\}$ *with* H *finite and* $K \bigcap H = \emptyset$ *we have* $\sum_{k \in H} x_i < \varepsilon$

Proof. First we have that If $s \in \{\sum_{i=k}^{n} x_i | n \in \{k, \ldots, \infty\}\}$ then $\exists n \in \{k, \ldots, \infty\}$ so that $s = \sum_{i=k}^{n} x_i = \sum_{i \in \{k, \ldots, n\}} x_i \in \{\sum_{i \in K} x_i | K \subseteq \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\}$ proving that

$$\left\{\sum_{i=k}^{n} x_i | n \in \{k, \ldots, \infty\}\right\} \in \left\{\sum_{i \in K} x_i | K \subseteq \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\right\} \tag{12.43}$$

1 ⇔ 2. Given $n \in \{k, \ldots, \infty\}$ we have that $\sum_{i=k}^{n+1} x_i = x_{n+1} + \sum_{i=k}^{n} x_i \geqslant \sum_{i=k}^{n} x_i$ proves that $\{\sum_{i=k}^{n}\}_{n \in \{k, \ldots, \infty\}}$ is a increasing sequence. the previous lemma (see 12.354) proves then that $1 \Leftrightarrow 2$

2 ⇒ 3. Take $s \in \{\sum_{n \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\} \wedge K \text{ is finite})\}$ then there exists a finite $K \subseteq \{k, \ldots, \infty\}$ so that $s = \sum_{i \in K} x_i$. Take then $n = \max(K)$ then as $i \in K \Rightarrow k \leqslant i \leqslant n = \{k, \ldots, n\}$ proving that $K \subseteq \{k, \ldots, n\}$. So $s = \sum_{i \in K} x_i \leqslant \sum_{i \in K} x_i + \sum_{i \in \{1, \ldots, n\} \setminus K} x_i = \sum_{i \in \{1, \ldots, n\}} x_i = \sum_{i=1}^{n} x_i \leqslant \sup(\{\sum_{i=k}^{n} x_i | n \in \{k, \ldots, \infty\}\})$ proving by 9.43 that

$$\sup\left(\left\{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\right\}\right) \text{exists} \tag{12.44}$$

exists and that

$$\sup\left(\left\{\sum_{i\in K} x_i | K\in\mathcal{P}(\{k,\ldots,\infty\})\wedge K \text{ is finite}\right\}\right)\leqslant\sup\left(\left\{\sum_{i=k}^{n} x_i | n\in\{k,\ldots,\infty\}\right\}\right) \quad (12.45)$$

Using 12.43 we have that $\sup\left(\{\sum_{i=k}^{n} x_i | n\in\{k,\ldots,\infty\}\}\right)\leqslant\sup\left(\{\sum_{i\in K} x_i | K\subseteq\mathcal{P}(\{k,\ldots,\infty\})\wedge K \text{ is finite}\}\right)$ which together with 12.45 proves that

$$\sup\left(\left\{\sum_{i\in K} x_i | K\in\mathcal{P}(\{k,\ldots,\infty\})\wedge K \text{ is finite}\right\}\right)=\sup\left(\left\{\sum_{i=k}^{n} x_i | n\in\{k,\ldots,\infty\}\right\}\right)$$

3 ⇒ 2. Let $s\in\{\sum_{i=k}^{n} x_i | n\in\{k,\ldots,\infty\}\}$ then from 12.43 we have that $s\in\{\sum_{i\in K} x_i | K\subseteq\mathcal{P}(\{k,\ldots,\infty\})\wedge K \text{ is finite}\}$ so that $s\leqslant\sup\left(\{\sum_{i\in K} x_i | K\subseteq\mathcal{P}(\{k,\ldots,\infty\})\wedge K \text{ is finite}\}\right)$ hence by 9.43 we have that

$$\sup\left(\left\{\sum_{i=k}^{n} x_i | n\in\{k,\ldots,\infty\}\right\}\right)\text{exists}$$

Applying then (2) proves the rest.

1 ⇒ 2. As $\sum_{i=k}^{\infty} x_i$ converges we have that $\lim_{n\rightarrow\infty}(\sum_{i=k}^{n})$ exists so that $\{\sum_{i=k}^{n}\}_{n\in\{k,\ldots,\infty\}}$ is Cauchy (see 12.311) hence there exists a $n\in\{k,\ldots,\infty\}$ such that $\forall r,s\geqslant n$ we have $|\sum_{i=k}^{r} x_i-\sum_{i=k}^{s} x_i|<\varepsilon$. Then if $l\geqslant n$ and $p\in\mathbb{N}$ we have $n\leqslant l\leqslant l+p$ so that $\sum_{i=l+1}^{l+p} x_i=\sum_{i=k}^{l+p} x_i-\sum_{i=k}^{l} x_i \underset{\sum_{i=k}^{l} x_i\leqslant\sum_{i=k}^{l+p} x_i}{=} |\sum_{i=k}^{l+p} x_i-\sum_{i=k}^{l} x_i|<\varepsilon$

2 ⇒ 1. Take $\varepsilon>0$ then there exists a $n\in\{k,\ldots,\infty\}$ such that $\forall l\geqslant n$ and $p\in\mathbb{N}$ we have that $\sum_{i=l+1}^{l+p} x_i<\varepsilon$. Take now $r,s\geqslant n$ then we may assume without loosing generality that $r\geqslant s$. If now $r=s$ then $|\sum_{i=k}^{r} x_i-\sum_{i=k}^{s}|=0<\varepsilon$ and if $s<r$ then $|\sum_{i=k}^{r} x_i-\sum_{i=k}^{s} x_i|=|\sum_{i=s+1}^{r} x_i|=\sum_{i=s+1}^{r} x_i \underset{0<p=r-s}{=} \sum_{i=s+1}^{s+p} x_i<\varepsilon$ proving that $\{\sum_{i=k}^{n} x_i\}_{n\in\{k,\ldots,\infty\}}$ is Cauchy. As $\mathbb{R}$ is complete $\{\sum_{i=k}^{n} x_i\}_{n\in\{k,\ldots,\infty\}}$ converges hence $\sum_{i=k}^{\infty} x_i$ converges

4 ⇒ 5. Let $\varepsilon>0$ then there exists a $n\in\{k,\ldots,\infty\}$ such that $\forall p\in\mathbb{N}$ and $l\geqslant n$ we have $\sum_{i=l+1}^{l+p} x_i<\varepsilon$. Take $K=\{k,\ldots,n\}$ then if $H\subseteq\{k,\ldots,\infty\}$ is finite with $H\bigcap K=\emptyset$ then we have either

***K* = ∅.** then $\sum_{i\in K} x_i=0<\varepsilon$

$K \neq \emptyset$. then $n < \min(K) \leqslant \max(K)$ so that $K \subseteq \{n+1, \ldots, n+p\}$ where $p = \max(K) - n \in \mathbb{N}$ so that $\sum_{i \in K} x_i \leqslant \sum_{i \in K} x_i + \sum_{i \in \{n+1, \ldots, n+p\}} x_i = \sum_{i=n+1}^{n+p} < \varepsilon$

proving in all cases $\sum_{i \in K} x_i < \varepsilon$

5 ⇒ 4. Let $\varepsilon > 0$ then there exists a finite $K \subseteq \{k, \ldots, \infty\}$ such that $\forall H \subseteq \{k, \ldots, \infty\}$ with H finite and $K \bigcap H = \emptyset$ we have $\sum_{k \in H} x_i < \varepsilon$. Take $K' = K \bigcup \{k\}$ then $\emptyset \neq K'$ s0 that $n = \max(K')$ exists. If now $n \leqslant l$ and $p \in \mathbb{N}$ then $K \bigcap \{l+1, \ldots, l+p\} \subseteq K' \bigcap \{l+1, \ldots, l+p\} = \emptyset$ hence $\sum_{i=l+1}^{l+p} x_i = \sum_{i \in \{l+1, \ldots, l+p\}} x_i < \varepsilon$ □

Corollary 12.371. *Let $\sum_{i=k}^{\infty} x_i$ be convergent serie such that $\forall i \in \{k, \ldots, \infty\}$ we have $0 < x_i$ then $0 < \sum_{i=k}^{\infty} x_i$*

Proof. $0 < x_k \leqslant \sup(\{\sum_{i=k}^{n} x_i | n \in \{k, \ldots, \infty\}\}) = \sum_{i=k}^{\infty} x_i$ □

The above equivalences can be used to prove the commutativity of series of positive numbers

Theorem 12.372. *Let $\{x_i\}_{i \in \{k, \ldots, \infty\}} \subseteq [0, \ldots, \infty[$ and $\beta: \{k, \ldots, \infty\} \rightarrow \{k, \ldots, \infty\}$ a bijection then if $\sum_{i=k}^{\infty} x_i$ exists we have that $\sum_{i=k}^{\infty} x_{\beta(i)}$ exists and $\sum_{i=k}^{\infty} x_i = \sum_{i=k}^{\infty} x_{\beta(i)}$*

Proof. If $s \in \{\sum_{i \in K} x_i | K \subseteq \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\}$ then there exists a $K \subseteq \{k, \ldots, \infty\}$ such that $s = \sum_{i \in K} x_i$. As $(\beta^{-1})_{|\beta(K)}: \beta(K) \rightarrow K$ is a bijection we have using 10.44 that $\sum_{i \in K} x_i = \sum_{i \in \beta^{-1}(K)} x_{\beta(i)} \in \{\sum_{i \in K} x_{\beta(i)} | K \subseteq \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\}$ [As $\beta^{-1}(K) \subseteq \{k, \ldots, \infty\}$ is finite] so we have that

$$\left\{\sum_{i \in K} x_i | K \subseteq \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\right\} \subseteq \left\{\sum_{i \in K} x_{\beta(i)} | K \subseteq \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\right\} \tag{12.46}$$

If $s \in \{\sum_{i \in K} x_{\beta(i)} | K \subseteq \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\}$ then there exists a $K \subseteq \{k, \ldots, \infty\}$ such that $s = \sum_{i \in K} x_{\beta(i)}$ then as $\beta_K: K \rightarrow \beta(K)$ is a bijection we can use 10.44 again giving $\sum_{i \in K} x_{\beta(i)} = \sum_{i \in \beta(K)} x_i \in \{\sum_{i \in K} x_i | K \subseteq \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\}$[as $\beta(K) \subseteq \{k, \ldots, \infty\}$ is finite). So we have proved that $\{\sum_{i \in K} x_{\beta(i)} | K \subseteq \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\} \subseteq \{\sum_{i \in K} x_i | K \subseteq \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\}$ which using 12.46 gives

$$\left\{\sum_{i \in K} x_i | K \subseteq \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\right\} \subseteq \left\{\sum_{i \in K} x_{\beta(i)} | K \subseteq \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\right\} \tag{12.47}$$

As $x = \sum_{i=k}^{\infty} x_i$ exists we have using 12.370 that $x = \sup(\{\sum_{i \in K} x_i | K \subseteq \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\}) \underset{12.47}{=} \sup(\{\sum_{i \in K} x_{\beta(i)} | K \subseteq \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\}) = \sum_{i=k}^{\infty} x_{\beta(i)}$ □

The above theorem allows us to extend the definition of a serie of positive numbers. Later we wil use this extension also for absolute convergent series.

Definition 12.373. *Let I be a countable set, $\{x_i\}_{i\in I}\subseteq[0,\infty[$ a countable family of positive numbers then we define convergence of $\sum_{i\in I}x_i$ as follows*

1. *If I is finite then $\sum_{i\in I}x_i$ converges and $\sum_{i\in I}x_i$ is as defined in 10.37). Using 10.40 we have then that $\sum_{i\in I}x_i=\sum_{i=0}^{n}x_{\beta(i)}$ where $\beta:\{0,\ldots,n\}\rightarrow I$ is a bijection.*
2. *If I is infinite then $\sum_{i\in I}x_i$ converges if $\sum_{i=0}^{\infty}x_{\beta(i)}$ converges where $\beta:\mathbb{N}_0\rightarrow I$ is a bijection (note that as I is inifinte and countable such a bijection exists). We denote then $\sum_{i\in I}x_i$ by $\sum_{i=0}^{\infty}x_{\beta(i)}$*

we must of course prove that this definition is independent of the choice of β

Proof.

1. This follows from 10.40
2. Let $\alpha:\mathbb{N}_0\rightarrow I$ be a another bijection then $\alpha^{-1}\circ\beta:\mathbb{N}_0\rightarrow\mathbb{N}_0$ is a bijection. So using 12.372 we have that $\sum_{i=0}^{\infty}x_{\alpha(\alpha^{-1}(\beta(i)))}$ is convergent and $\sum_{i=0}^{\infty}x_{\alpha(i)}=\sum_{i=0}^{\infty}x_{\alpha(\alpha^{-1}(\beta(i)))}=\sum_{i=0}^{\infty}x_{\beta(i)}=s$ □

Lemma 12.374. *Let I,J be countable sets, $I\subseteq J$ and $\{x_i\}_{i\in J}\subseteq[0,\infty[$ such that $\sum_{i\in J}x_i$ exists then we have that $\sum_{i\in I}x_i$ and $\sum_{i\in I}x_i\leqslant\sum_{i\in J}x_i$*

Proof. We must consider the following cases for J

***J* is finite.** then as $I\subseteq J$ we have that I is also finite and $\sum_{i\in J}x_i\underset{10.46}{=}\sum_{i\in J\setminus I}x_i+\sum_{i\in I}x_i\geqslant\sum_{i\in I}x_i$

***J* is infinite.** By definition there exists a bijection $\beta:\mathbb{N}_0\rightarrow J$ such that $\sum_{i\in J}x_i=\sum_{i=0}^{\infty}x_{\beta(i)}$. For I we have either

***I* is finite.** As I is finite and β is a bijection we have that $\beta^{-1}(I)\subseteq\mathbb{N}_0$ is finite so that $\sum_{i\in I}x_i\underset{10.44}{=}\sum_{i\in\beta^{-1}(I)}x_{\beta(i)}\leqslant\sup(\{\sum_{i\in K}x_{\varphi(i)}|K\subseteq\mathbb{N}_0\wedge K\text{ is finite}\})=\sum_{i=0}^{\infty}x_{\varphi(i)}=\sum_{i\in J}x_i$ proving that $\sum_{i=I}x_i\leqslant\sum_{i\in J}x_i$

***J* is infinite.** Then there exists a bijection $\alpha:\mathbb{N}_0\rightarrow I\subseteq J$. Take a finite $K\subseteq\mathbb{N}_0$ then $\alpha(K)\subseteq J\wedge\alpha(K)$ is finite so that $\beta^{-1}(\alpha(K))$ is finite. As we have that

$$\begin{aligned}\sum_{i\in K}x_{\alpha(i)} &\underset{10.44}{=} \sum_{i\in\alpha(K)}x_i\\ &\underset{10.44}{=} \sum_{i\in\beta^{-1}(\alpha(K))}x_{\beta(i)}\\ &\leqslant \sup\left(\left\{\sum_{i\in L}x_{\beta(i)}|L\subseteq\mathbb{N}_0\wedge L\text{ is finite}\right\}\right)\\ &\leqslant \sum_{i=0}^{\infty}x_{\beta(i)}\end{aligned}$$

Proving by 9.43 that $\sup\left(\left\{\sum_{i\in K} x_{\alpha(i)}|K\subseteq N\wedge K\text{ is finite}\right\}\right)$ exists and $\sup\left(\left\{\sum_{i\in K} x_{\alpha(i)}|K\subseteq N\wedge K\text{ is finite}\right\}\right)\leqslant\sum_{i=0}^{\infty} x_{\beta(i)}$ proving that $\sum_{i=-}^{\infty} x_{\alpha(i)}$ exists and $\sum_{i=0}^{\infty} x_{\alpha(i)}\leqslant\sum_{i=0}^{\infty} x_{\beta(i)}$. So by definition we have that $\sum_{i\in I} x_i$ exists and $\sum_{i\in I} x_i\leqslant\sum_{i\in J} x_i$. □

Lemma 12.375. *Let I be countable set and $\{x_i\}_{i\in I}\subseteq[0,\infty[$ then we have the following equivalences*

1. $\sum_{i\in I} x_i$ *converges to s*
2. $\sup\left(\left\{\sum_{i\in K} x_i|K\subseteq I\wedge K\text{ is finite}\right\}\right)$ *exists and* $\sup\left(\left\{\sum_{i\in K} x_i|K\subseteq I\wedge K\text{ is finite}\right\}\right)=s$

Proof. *As I is countable we have two possibilities*

***I* is finite.** *then $\sum_{i\in I} x_i$ converges by definition and as $\forall K\subseteq I$ we have K is finite and* $\sum_{i\in K} x_i\leqslant\sum_{i\in K} x_i+\sum_{i\in I\setminus K} x_i\underset{10.46}{=}\sum_{i\in I} x_i$ *so that* $\sum_{i\in I} x_i=\max\left(\left\{\sum_{i\in K} x_i|K\subseteq I\wedge K\text{ is finite}\right\}\right)=\sup\left(\left\{\sum_{i\in K} x_i|K\subseteq I\wedge K\text{ is finite}\right\}\right)$ *(which of course exists)*

***I* is infinite.** *Then there exists a bijection $\beta:\mathbb{N}_0\to I$. If now $s\in\left\{\sum_{i\in K} x_{\beta(i)}|K\subseteq\mathbb{N}_0\wedge K\text{ is finite}\right\}$ then $s=\sum_{i\in K} x_{\beta(i)}\underset{10.44}{=}\sum_{i\in\beta(K)} x_i\in\left\{\sum_{i\in K} x_i|K\subseteq I\wedge K\text{ is finite}\right\}$, also if $s\in\left\{\sum_{i\in K} x_i|K\subseteq I\wedge K\text{ is finite}\right\}$ then $s=\sum_{i\in K} x_i\underset{10.44}{=}\sum_{i\in\beta^{-1}(K)} x_{\beta(i)}\in\left\{\sum_{i\in K} x_{\beta(i)}|K\subseteq\mathbb{N}_0\wedge K\text{ is finite}\right\}$ proving that*

$$\left\{\sum_{i\in K} x_{\beta(i)}|K\subseteq\mathbb{N}_0\wedge K\text{ is finite}\right\}=\left\{\sum_{i\in K} x_i|K\subseteq I\wedge K\text{ is finite}\right\}\tag{12.48}$$

Now

$$\begin{aligned}\sum_{i\in I} x_i\textit{ converges to } s &\underset{\text{definition}}{\Leftrightarrow} \sum_{i=0}^{\infty} x_{\beta(i)}\textit{ converges to } s\\ &\underset{12.370}{\Leftrightarrow} \sup\left(\left\{\sum_{i\in K} x_{\beta(i)}|K\subseteq\mathbb{N}_0\wedge K\text{ is finite}\right\}\right)\\ &\qquad\textit{exists and is equal to } s\\ &\underset{12.48}{\Leftrightarrow} \sup\left(\left\{\sum_{i\in K} x_i|K\subseteq I\wedge K\text{ is finite}\right\}\right)\textit{ exists}\\ &\qquad\textit{and is equal to } s\end{aligned}$$

□

Example 12.376. Let I be a countable set and $\{x_i\}_{i\in I}$ be such that $\forall i\in I$ we have $x_i=0$ then $\sum_{i\in I} x_i$ converges to 0

Proof. If $K \subseteq I$ finite then we have as $\forall i \in K$ $x_i = 0$ that $\sum_{i \in K} x_i \underset{10.38}{=} 0$ so that $\sup(\{\sum_{i \in K} x_i | K \subseteq I \wedge K \text{ is finite}\}) = 0$ which by the previous lemma (see 12.375) proves the example. □

Next we proceed to prove that series of positive numers are associative. First we need a little lemma concerning supremums in $\mathbb{R}$.

Lemma 12.377. *Let $n \in \mathbb{N}_0$ then if $\{A_i\}_{i \in \{0,\ldots,n\}}$ be such that $\forall i \in \{0,\ldots,n\} \vDash A_i \subseteq \mathbb{R}$ and $\exists s \in \mathbb{R}$ so that $\forall (x_0, \ldots, x_n) \in \prod_{i \in \{0,\ldots,n\}} A_i$ we have $\sum_{i=0}^n x_i \leqslant s$ then $\forall i \in \{0,\ldots,n\}$ $\sup(A_i)$ exists and $\sum_{i=0}^n \sup(A_i) \leqslant s$*

Proof. We prove this by induction so let $\mathcal{S} = \{n \in \mathbb{N}_0 | \text{if } \{A_i\}_{i \in \{0,\ldots,n\}} \subseteq \mathbb{R} \text{ be such that } \exists s \in \mathbb{R} \text{ so that } \forall (x_0, \ldots, x_n) \in \prod_{i \in \{0,\ldots,n\}} A_i \text{ we have } \sum_{i=0}^n x_i \leqslant s \text{ then } \forall i \in \{0,\ldots,n\} \sup(A_i) \text{ exists and } \sum_{i=0}^n \sup(A_i) \leqslant s\}$ then

- **$0 \in \mathcal{S}$.** Let $(x_0) \in \prod_{i=\{0,\ldots,0\}} A_i$ then $x_0 = \sum_{i=0}^0 x_i \leqslant s$ so that $\sup(A_0)$ exists and $\sum_{i=0}^0 \sup(A_0) = \sup(A_0) \leqslant s$
- **$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$.** Let $\{A_i\}_{i \in \{0,\ldots,n+1\}} \subseteq \mathbb{R}$ be such that $\exists s \in \mathbb{R}$ so that $\forall (x_0, \ldots, x_{n+1}) \in \prod_{i \in \{0,\ldots,n+1\}} A_i$ we have $\sum_{i=0}^{n+1} x_i \leqslant s$. Take then $x \in A_{n+1}$ and $(x_0, \ldots, x_n) \in \prod_{i \in \{0,\ldots,n\}} A_i$ then we have as $(x_0, \ldots, x_n, x) \in \prod_{i \in \{0,\ldots,n+1\}} A_i$ that $\sum_{i=0}^n x_i + x = \sum_{i=0}^{n+1} x_i \leqslant s$ so that $x \leqslant s - \sum_{i=0}^n x_i$ hence by 9.43 we have that $\sup(A_{n+1})$ exists and $\sup(A_{n+1}) \leqslant s - \sum_{i=0}^n x_i \Rightarrow \sum_{i=0}^n x_i \leqslant s - \sup(A_{n+1})$. As $n \in \mathcal{S}$ we have then that $\sum_{i=0}^n \sup(A_i) \leqslant s - \sup(A_{n+1}) \Rightarrow \sum_{i=0}^n \sup(A_i) + \sup(A_{n+1}) \leqslant s$ proving that $\sum_{i=0}^{n+1} \sup(A_i) \leqslant s$ and hence $n+1 \in \mathcal{S}$ □

Lemma 12.378. *Let I be a countable set, $\sum_{i \in I} x_i$ a convergent serie with $\{x_i\}_{i \in I} \subseteq [0, \infty[$, $n \in \mathbb{N}_0$ and $\{N_i\}_{i \in \{0,\ldots,n\}}$ such that $\forall i, j \in \{0,\ldots,n\} \vdash i \neq j$ we have $N_i \bigcap N_j = \emptyset$ and $I = \bigcup_{i \in \{0,\ldots,n\}} N_i$ then $\forall i \in \{0,\ldots,n\}$ we have that $\forall i \in \{0,\ldots,n\}$ $\sum_{j \in N_i} x_i$ converges and $\sum_{i \in I} x_i = \sum_{i=0}^n (\sum_{j \in N_i} x_j)$*

Proof. First as $\forall i \in \{0,\ldots,n\}$ $N_i \subseteq I$ it follows from the convergence of $\sum_{i \in I} x_i$ and lemma 12.374 that

$$\forall i \in \{0,\ldots,n\} \vDash \sum_{i \in N_i} x_i \text{ converges} \tag{12.49}$$

Second as $\sum_{i \in I} x_i \underset{12.375}{=} \sup(\{\sum_{i \in K} x_i | K \subseteq I \wedge K \text{ finite}\})$ we have given $\varepsilon > 0$ that there exists a $K \subseteq I$ finite so that $\sum_{i \in I} x_i - \varepsilon < \sum_{i \in K} x_i$ hence

$$\sum_{i \in I} x_i < \sum_{i \in K} x_i + \varepsilon \tag{12.50}$$

Now given $i\in\{0,\ldots,n\}$ we have that $N_i\bigcap K$ is finite, $K=I\bigcap K=K\bigcap\left(\bigcup_{i\in\{0,\ldots,n\}}N_i\right)=\bigcup_{i\in\{0,\ldots,n\}}(N_i\bigcap K)$ and $\forall i,j\in\{0,\ldots,n\}$ we have $(N_i\bigcap K)\bigcap(N_j\bigcap K)=(N_i\bigcap N_j)\bigcap K=\emptyset$. So

$$\begin{aligned}\sum_{i\in K}x_i &\underset{10.46}{=} \sum_{i=0}^{n}\left(\sum_{j\in N_i\bigcap K}x_j\right)\\ &\leqslant \sum_{i=0}^{n}\left(\sum_{j\in N_i}x_j\right)\left(\text{see }12.374, 12.49\text{ and }N_i\bigcap K\subseteq N_i\right)\end{aligned}$$

Using 12.50 together with the above gives then that $\sum_{i\in I}x_i<\sum_{i=0}^{n}\left(\sum_{j\in N_i}x_j\right)+\varepsilon$ which as $\varepsilon>0$ was arbitrary choosen means by 9.56 that

$$\sum_{i\in I}x_i\leqslant\sum_{i=0}^{n}\left(\sum_{j\in N_i}x_j\right)\tag{12.51}$$

Take now $(s_0,\ldots,s_n)\in\prod_{i\in\{0,\ldots,n\}}\left\{\sum_{i\in K}|K\subseteq N_i\wedge K\text{ finite}\right\}$ then there exists a family $\{K_i\}_{i\in\{0,\ldots,n\}}$ of sets with $\forall i\in\{0,\ldots,n\}$ $K_i\subseteq N_i$ (so the sets are finite and pairwise disjoint) and $s_i=\sum_{j\in K_i}x_j$. Hence $\sum_{i=0}^{n}s_i=\sum_{i=0}^{n}\left(\sum_{j\in K_i}x_j\right)\underset{10.46}{=}\sum_{j\in\bigcup_{i\in\{0,\ldots,n\}}K_i}x_i\leqslant\sum_{i\in I}x_i$ (as $\bigcup_{i\in\{0,\ldots,n\}}K_i\subseteq\bigcup_{i\in\{0,\ldots,n\}}N_i=I$ together with 12.374). Hence using 12.377 we have then $\sum_{i=0}^{n}\sup\left(\left\{\sum_{j\in K}x_j|K\subseteq N_i\right\}\right)\leqslant\sum_{i\in I}x_i$, so using 12.375 we have that

$$\sum_{i=0}^{n}\left(\sum_{j\in N_i}x_j\right)\leqslant\sum_{i\in I}x_i\tag{12.52}$$

Using 12.51 and 12.52 we have

$$\sum_{i\in I}x_i=\sum_{i=0}^{n}\left(\sum_{j\in N_i}x_j\right)$$

□

Lemma 12.379. *Let I be a countable set, $\sum_{i\in I}x_i$ a convergent serie with $\{x_i\}_{i\in I}\subseteq[0,\infty[$, $\{N_i\}_{i\in\mathbb{N}_0}$ be such that $\mathbb{N}_0=\bigcup_{i\in\mathbb{N}_0}N_i$ and $\forall i,j\in\mathbb{N}_0\vdash i\neq j$ $N_i\bigcap N_j=\emptyset$ then we have that $\forall i\in\mathbb{N}_0$ $\sum_{j\in N_i}x_j$ converges and $\sum_{i=0}^{\infty}\left(\sum_{j\in N_i}x_j\right)$ converges to $\sum_{i\in\mathbb{N}_0}x_i$*

Proof. First using 12.374 we have as $\forall i\in\mathbb{N}_0$ $N_i\subseteq I$ that

$$\forall i\in\mathbb{N}_0\vDash\sum_{j\in N_i}x_j\text{ converges}\tag{12.53}$$

Let $n \in \mathbb{N}_0$ then $K_n = \bigcup_{i \in \{0,\ldots,n\}} N_i \subseteq I$ so by 12.374 we have that $\sum_{i \in K_n} x_i$ converges and $\sum_{i \in K_n} x_i \leqslant \sum_{i \in I} x_i$. Further by 5.66 we have that K_n is countable and as $\{N_i\}_{i \in \{0,\ldots,n\}}$ is pairwise disjoint we can use the previous lemma (see 12.378) giving $\sum_{i=0}^{n} (\sum_{j \in N_i} x_j) = \sum_{i \in K_n} x_i \leqslant \sum_{i \in I} x_i$. As this is true for every $n \in \mathbb{N}_0$ we have by 9.43 that $\sup(\{\sum_{i=0}^{n} (\sum_{j \in N_i} x_j) | n \in \mathbb{N}_0\})$ exists and $\sup(\{\sum_{i=0}^{n} (\sum_{j \in N_i} x_j) | n \in \mathbb{N}_0\}) \leqslant \sum_{i \in I} x_i$. Using then 12.370 we conclude that

$$\sum_{i=0}^{\infty} \left(\sum_{j \in N_i} x_j \right) \text{converges and} \sum_{i=0}^{\infty} \left(\sum_{j \in N_i} x_j \right) \leqslant \sum_{i \in I} x_i \tag{12.54}$$

For the opposite equility, let $\varepsilon > 0$ then as by 12.375 we have that $\sum_{i \in I} x_i = \sup(\{\sum_{i \in K} x_i | K \subseteq I \wedge K \text{ is finite}\})$ there exists a finite K such that $\sum_{i \in I} x_i - \varepsilon < \sum_{i \in K} x_i$ hence

$$\sum_{i \in I} x_i < \sum_{i \in K} x_i + \varepsilon \tag{12.55}$$

Take now $J = \{i \in \mathbb{N}_0 | K \bigcap N_i \neq \emptyset\} \subseteq I$ so that by 5.66 J is countable. Then $\bigcup_{i \in J} (K \bigcap N_i) \subseteq K$ and if $x \in K = K \bigcap \mathbb{N}_0 = K \bigcap (\bigcup_{i \in \mathbb{N}_0} N_i) = \bigcup_{i \in \mathbb{N}_0} (K \bigcap N_i)$ there exists a $i \in \mathbb{N}_0$ such that $x \in K \bigcap N_i \Rightarrow K \bigcap N_i \neq \emptyset$ hence $x \in K \bigcap N_i$ where $i \in I$. So we have that

$$K = \bigcup_{i \in J} (K \bigcap N_i) \tag{12.56}$$

Now if J is infinite then as $\{K \bigcap N_i\}_{i \in J}$ is a family of pairwise disjoint non empty sets we would by 5.72 have that K is infinite a contradiction. So J is finite, then we have either

***$J = \emptyset$*.** then if $\forall i \in \mathbb{N}_0$ we have $K \bigcap N_i = \emptyset \Rightarrow \emptyset = \bigcup_{i \in \mathbb{N}_0} (K \bigcap N_i) = K \bigcap (\bigcup_{i \in \mathbb{N}_0} N_i) = K \bigcap \mathbb{N}_0 = K$ so that $\sum_{i \in I} x_i < 0 + \varepsilon$ which as ε is choosen arbitrary means $\sum_{i \in I} x_i \leqslant 0 \leqslant \sum_{i=0}^{\infty} (\sum_{j \in N_i} x_j)$

***$J \neq \emptyset$*.** then as J is finite $m = \max(J)$ exists. If now $x \in \bigcup_{i \in \{0,\ldots,m\}} (K \bigcap N_i)$ then $\exists i \in \{0,\ldots,m\}$ so that $x \in K \bigcap N_i \Rightarrow K \bigcap N_i \neq \emptyset \Rightarrow i \in J \Rightarrow x \in \bigcup_{i \in J} (K \bigcap N_i)$, also if $x \in \bigcup_{i \in J} (K \bigcap N_i)$ then $\exists i \in J \subseteq \{0,\ldots,m\}$ such that $x \in K \bigcap N_i \Rightarrow x \in \bigcup_{i \in \{0,\ldots,m\}} (K \bigcap N_i)$. So we have

$$K = \bigcup_{i \in J} (K \bigcap N_i) = \bigcup_{i \in \{0,\ldots,m\}} (K \bigcap N_i) \tag{12.57}$$

As $\{K \bigcap N_i\}_{i \in \{0,\ldots,m\}}$ is pairwise disjoint we have using the above and the previous lemma (see 12.378) that $\sum_{i \in K} x_i \underset{12.378}{=} \sum_{i=0}^{m} (\sum_{j \in (K \bigcap N_i)} x_j) \leqslant_{12.374} \sum_{i=0}^{m} (\sum_{j \in N_i} x_j) \leqslant_{12.370} \sum_{i=0}^{\infty} (\sum_{j \in N_i} x_j)$. Hence by 12.55 we have that $\sum_{i \in I} x_i < \sum_{i=0}^{\infty} (\sum_{j \in N_i} x_j) + \varepsilon$ proving that $\sum_{i \in I} x_i \leqslant \sum_{i=0}^{\infty} (\sum_{j \in N_i} x_j)$

As we have in all cases that $\sum_{i\in I} x_i \leqslant \sum_{i=0}^{\infty} (\sum_{j\in N_i} x_j)$ we can use 12.54 to prove that

$$\sum_{i\in I} x_i = \sum_{i=0}^{\infty} \left(\sum_{j\in N_i} x_j \right)$$ □

We can easely extend the above to the following general case of countable families

Theorem 12.380. *Let I, K be countable sets, $I \neq \emptyset$, $\{K_i\}_{i\in I}$ a family of countable sets such that $K = \bigcup_{i\in I} K_i$ and $\forall i, j \in I$ $K_i \bigcap K_j = \emptyset$ then for every $\{x_i\}_{i\in I} \subseteq [0, \infty[$ such that $\sum_{i\in K} x_i$ converges we have that $\forall i \in I$ $\sum_{j\in K_i} x_i$ converges and $\sum_{i\in I} (\sum_{j\in K_i} x_j)$ converges to $\sum_{i\in K} x_i$*

Proof. As I is countable we have two possible cases to consider

***I* is finite.** then there exists a $n \in \mathbb{N}_0$ and a bijection $\beta: \{0, \ldots, n\} \to I$. Define now $\{L_i\}_{i\in\{0,\ldots,n\}}$ by $L_i = K_{\beta(i)}$ then we have $\bigcup_{i\in\{0,\ldots,n\}} L_i = \bigcup_{i\in\{0,\ldots,n\}} K_{\beta(i)} \underset{2.64}{=} \bigcup_{i\in I} K_i = K$, further $\forall i, j \in \{0, \ldots, n\}$ with $i \neq j$ we have as β is a bijection that $\beta(i) = \beta(j)$ so that $L_i \bigcap L_j = K_{\beta(i)} \bigcap K_{\beta(j)} = \emptyset$. So we can use 12.377 giving that $\forall i \in \{0, \ldots, n\}$ $\sum_{j\in L_i} x_i$ converges and $\sum_{i\in\{0,\ldots,n\}} \left(\sum_{j\in L_j} x_j\right) = \sum_{i\in K} x_i$. Finally $\forall i \in I$ we have that $K_i = K_{\beta(\beta^{-1}(i))} = L_{\beta^{-1}(i)}$ so that $\sum_{j\in K_i} x_j$ converges and $\sum_{i\in I} (\sum_{j\in K_i} x_j) \underset{\text{definition (see 12.373)}}{=} \sum_{i\in\{0,\ldots,n\}} (\sum_{j\in K_i} x_j) = \sum_{i\in K} x_i$.

***I* is infinite.** there exists a bijection $\beta: \mathbb{N}_0 \to I$. Define now $\{L_i\}_{i\in\mathbb{N}_0}$ by $L_i = K_{\beta(i)}$ then we have $\bigcup_{i\in\mathbb{N}_0} L_i = \bigcup_{i\in\mathbb{N}_0} K_{\beta(i)} \underset{2.64}{=} \bigcup_{i\in I} K_i = K$, further $\forall i, j \in \mathbb{N}_0$ with $i \neq j$ we have as β is a bijection that $\beta(i) = \beta(j)$ so that $L_i \bigcap L_j = K_{\beta(i)} \bigcap K_{\beta(j)} = \emptyset$. So we can use the previous lemma (see 12.379) which gives that $\forall i \in \mathbb{N}_0$ we have $\sum_{j\in L_i} x_j$ converges and $\sum_{i=0}^{\infty} \left(\sum_{j\in L_j} x_j\right)$ converges to $\sum_{i\in K} x_i$. Finally $\forall i \in I$ we have that $K_i = K_{\beta(\beta^{-1}(i))} = L_{\beta^{-1}(i)}$ so that $\sum_{j\in K_i} x_j$ converges and $\sum_{i\in I} (\sum_{j\in K_i} x_j) \underset{\text{definition (see 12.373)}}{=} \sum_{i=0}^{\infty} (\sum_{j\in K_i} x_j) = \sum_{i\in K} x_i$. □

12.11.7 Absolute convergent series

Theorem 12.381. (Dominent Convergence) *Let $\langle X, \|\|\rangle$ be a Banach space $\sum_{i=k}^{\infty} x_i$ a serie such that there exists a convergent serie $\sum_{i=k}^{\infty} s_i$ such that $\forall i \in \{k, \ldots, \infty\}$ we have $s_i \in \mathbb{R} \wedge \|x_i\| \leqslant s_i$ (so $s_i \geqslant 0$) then $\sum_{i=k}^{\infty} x_i$ converges and $\|\sum_{i=k}^{\infty} x_i\| \leqslant \sum_{i=k}^{\infty} s_i$*

Proof. First we prove that $\{\sum_{i=k}^{n} x_i\}_{n\in\{k,\ldots,\infty\}}$ is Cauchy. So let $\varepsilon > 0$ then as $\{\sum_{i=k}^{n}\}_{n\in\{k,\ldots,\infty\}}$ converges it is Cauchy and thus $\exists N$ such that if $n, m \geqslant N$ then $|\sum_{i=k}^{n} s_i - \sum_{i=k}^{m} s_i| < \varepsilon$. Consider now the following sub cases for n, m:

1. $(\boldsymbol{n = m})$ then $\|\sum_{i=k}^{n} x_i - \sum_{i=k}^{m} x_i\| = 0 < \varepsilon$
2. $(\boldsymbol{n < m})$ then $\|\sum_{i=k}^{n} x_i - \sum_{i=k}^{m} x_i\| = \|\sum_{i=k}^{m} x_i - \sum_{i=k}^{n} x_i\| = \|\sum_{i=n+1}^{m} x_i\| \leqslant_{12.70} \sum_{i=n+1}^{m} \|x_i\| \leqslant_{10.63} \sum_{i=n+1}^{m} |s_i| \underset{s_i \geqslant 0}{=} \sum_{i=n+1}^{m} s_i \underset{0 \leqslant \sum_{i=n+1}^{m} s_i}{=} |\sum_{i=n+1}^{m} s_i| = |\sum_{i=k}^{m} s_i - \sum_{i=k}^{n} s_i| = |\sum_{i=k}^{n} s_i - \sum_{i=k}^{m} s_i| < \varepsilon$
3. $(\boldsymbol{n > m})$ then $\|\sum_{i=k}^{n} x_i - \sum_{i=k}^{m} x_i\| = \|\sum_{i=m+1}^{n} x_i\| \leqslant_{12.70} \sum_{i=m+1}^{n} \|x_i\| \leqslant_{10.63} \sum_{i=m+1}^{n} |s_i| = \sum_{i=m+1}^{n} s_i \underset{0 \leqslant \sum_{i=n+1}^{m} s_i}{=} |\sum_{i=m+1}^{n} s_i| = |\sum_{i=k}^{n} s_i - \sum_{i=k}^{m} s_i| < \varepsilon$ (as $s_i \geqslant 0 \Rightarrow |s_i| = s_i$) □

So in all cases we have $\|\sum_{i=k}^{n} x_i - \sum_{i=k}^{m} x_i\| < \varepsilon$ proving that $\{\sum_{i=k}^{n}\}_{n\in\{k,\ldots,\infty\}}$ is Cauchy. as $\langle X, \|\|\rangle$ is a Banach space and thus complete we have that $\{\sum_{i=k}^{n}\}_{n\in\{k,\ldots,\infty\}}$ converges and thus that $\sum_{i=k}^{\infty} x_i$ exist.

Second if $n \in \{k, \ldots\infty\}$ then we have $\|\sum_{i=k}^{n} x_i\| \leqslant \sum_{i=k}^{n} \|x_i\| \leqslant \sum_{i=k}^{n} |x_i| = \sum_{i=k}^{n} s_i$ so that by 12.339 we have that $\|\sum_{i=k}^{\infty} x_i\| = \left\|\lim_{n\to\infty} \sum_{i=k}^{n} x_i\right\| \leqslant_{12.338} \lim_{n\to\infty} \sum_{i=k}^{n} s_i$.

Corollary 12.382. *Let $\langle X, \|\|\rangle$ is a normed Banach space, $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ a sequence so that $\sum_{i=k}^{\infty} \|x_i\|$ converges then $\sum_{i=k}^{\infty} x_i$ converges and $\|\sum_{i=k}^{\infty} x_i\| \leqslant \sum_{i=k}^{\infty} \|x_i\|$*

Proof. This follows from the previous theorem as $\{\|x_i\|\}_{i\in\{k,\ldots,\infty\}}$ is a sequence in $\mathbb{R}$ and $\forall i \in \{k, \ldots, \infty\}$ we have $\|x_i\| \leqslant \|x_i\|$. □

the above theorem motivates the definition of absolute convergence of a serie.

Definition 12.383. (Absolute Convergence) *Let $\langle X, \|\|\rangle$ be a normed Banach space then $\sum_{i=k}^{\infty} x_i$ is **absolute convergent** if and only if $\sum_{i=k}^{\infty} \|x_i\|$ is convergent,*

Note 12.384. Using 12.382 we have that absolute convergence implies convergence.

We set now out to prove that absolute convergent series in a Banach space are commutative and associative. First we need some theorems and lemmas.

Theorem 12.385. *Let $\langle X, \|\|\rangle$ be a normed space then we have the following equivalences for $\sum_{i=k}^{\infty} x_i$*

1. *$\sum_{i=k}^{\infty} x_i$ is absolute convergent*
2. *$\forall \varepsilon > 0$ there exists a $n \in \{k, \ldots, \infty\}$ such that $\forall n \leqslant l$ and $p \in \mathbb{N}$ we have $\sum_{i=l+1}^{l+p} \|x_i\| < \varepsilon$*

3. *$\forall \varepsilon > 0$ there exists a finite $K \subseteq \{k, ..., \infty\}$ such that $\forall H \subseteq \{k, ..., \infty\}$ with H finite and $K \bigcap H = \emptyset$ we have $\sum_{i \in H} \|x_i\| < \varepsilon$*

Proof. As $\{\|x_i\|\}_{i \in \{k, ..., \infty\}}$ is a sequence of positive numbers and absolute convergence means that $\sum_{i=k}^{\infty} \|x_i\|$ converges we can use 12.370 to prove the theorem. □

Lemma 12.386. *Let $\langle X, \|\|\rangle$ be a normed space, $\varepsilon > 0$ and $\sum_{i=k}^{\infty} x_i$ a absolute convergent serie then there exists a finite $K \subseteq \{k, ..., \infty\}$ such that $\forall \varepsilon > 0$ we have*

1. *$\forall H \subseteq \{k, ..., \infty\}$ with H finite and $K \subseteq H$ we have $\|\sum_{i=k}^{\infty} x_i - \sum_{i \in H} x_i\| < \varepsilon$*
2. *$\forall L \subseteq \{k, ..., \infty\}$ with L finite and $K \bigcap L = \emptyset$ we have $\sum_{i \in L} \|x_i\| < \varepsilon$*

Proof. Let $\varepsilon > 0$. As absolute convergence implies convergence (see 12.384) there exists a $n \in \{k, ..., \infty\}$ so that

$$\forall l \geqslant n \text{ we have } \left\| \sum_{i=k}^{l} x_i - \sum_{i=k}^{\infty} \right\| < \frac{\varepsilon}{2} \tag{12.58}$$

Using absolute convergence and the previous theorem (see 12.385) there exists a finite $K' \subseteq \{k, ..., \infty\}$ such that $\forall L \subseteq \{k, ..., \infty\}$ with L finite and $K' \bigcap L = \emptyset$ we have $\sum_{i \in L} \|x_i\| < \frac{\varepsilon}{2}$. Take now $K = \{k, ..., n\} \bigcup K'$ then we have

$$\{k, ..., n\} \subseteq K \wedge \forall L \subseteq \{k, ..., \infty\} \vdash L \text{ finite and } K \bigcap L = \emptyset \text{ we have } \sum_{i \in L} \|x_i\| < \frac{\varepsilon}{2} < \varepsilon \tag{12.59}$$

Let now $H \subseteq \{k, ..., \infty\}$ with $K \subseteq H$ and H finite. Let $p = \max(H) \Rightarrow H \subseteq \{k, ..., p\}$ then

$$\begin{aligned}
\left\| \sum_{i \in H} x_i - \sum_{i=k}^{\infty} x_i \right\| &= \left\| \sum_{i \in \{k, ...p\}} x_i - \sum_{i \in \{k, ..., p\} \setminus H} x_i - \sum_{i=k}^{\infty} x_i \right\| \\
&= \left\| \sum_{i=k}^{p} x_i - \sum_{i \in \{k, ..., p\} \setminus H} x_i - \sum_{i=k}^{\infty} x_i \right\| \\
&\leqslant \left\| \sum_{i=k}^{p} x_i - \sum_{i=k}^{\infty} x_i \right\| + \left\| \sum_{i \in \{k, ..., p\} \setminus H} x_i \right\| \\
&< \left\| \sum_{i=k}^{p} x_i - \sum_{i=k}^{\infty} x_i \right\| + \frac{\varepsilon}{2} \text{as } K \bigcap \{k, ..., p\} \setminus H \subseteq \{k, ..., p\} \setminus K = \emptyset \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \text{ as } \{k, ..., n\} \subseteq K \subseteq H \Rightarrow n \leqslant \max(H) = p \text{ and } 12.58 \\
&= \varepsilon
\end{aligned}$$

which proves (1). (2) follows from 12.59. □

Theorem 12.387. *Let $\langle X, \|\|\rangle$ be a Banach space, $k, m, n \in \mathbb{N}_0 \vdash m \leqslant n$ and $\{x_{i,j}\}_{(i,j) \in \{k,\ldots,\infty\} \times \{m,\ldots,n\}}$ a family in X then if $\forall j \in \{m, \ldots, n\}$ we have that $\sum_{i=k}^{\infty} x_{i,j}$ exists then $\sum_{i=k}^{\infty} (\sum_{j=m}^{n} x_{i,j})$ exists and is equal to $\sum_{j=m}^{n} (\sum_{i=k}^{\infty} x_{i,j})$*

Proof. As $\forall j \in \{m, \ldots, n\}$ we have that $b_j = \sum_{i=k}^{\infty} x_{i,j}$ exists there is given a $\varepsilon > 0$ a $N_j \geqslant 0$ such that if $l \geqslant N_j$ then $\|\sum_{i=k}^{l} x_{i,j} - b_j\| < \frac{\varepsilon}{n-m+1}$. then we have that if $l \geqslant \max(\{b_j | j \in \{m, \ldots, n\}\})$ that $\|\sum_{i=k}^{l} (\sum_{j=m}^{n} x_{i,j}) - \sum_{j=m}^{n} b_j\| \underset{10.12}{=} \|\sum_{j=m}^{n} (\sum_{i=k}^{l} x_{i,j}) - \sum_{j=m}^{n} b_j\| = \|\sum_{j=m}^{n} (\sum_{i=k}^{l} x_{i,j} - b_j)\| \leqslant_{12.70} \sum_{j=m}^{n} \|\sum_{i=k}^{l} x_{i,j} - b_j\| < \sum_{j=m}^{n} \frac{\varepsilon}{m-n+1} < \varepsilon$ proving that $\sum_{i=k}^{\infty} (\sum_{j=m}^{n} x_{i,j}) = \sum_{j=m}^{n} (\sum_{i=k}^{\infty} x_{i,j})$. □

We are now ready to prove comunitativity for absolute convergent series, first we prove that permutations conserves convergence.

Theorem 12.388. *Let $\langle X, \|\|\rangle$ be a Banach space and $\sum_{i=k}^{\infty} x_i$ a absolute convergent serie then for every permutation $\beta: \{k, \ldots, \infty\} \rightarrow \{k, \ldots, \infty\}$ we have that $\sum_{i=k}^{\infty} x_{\beta(i)}$ is absolute convergent. If $\langle X, \|\|\rangle$ is a Banach space then it follows from 12.384 that $\sum_{i=k}^{\infty} x_{\beta(i)}$ is convergent.*

Proof. As $\sum_{i=k}^{\infty} x_i$ is Absolute convergent we have that for the sequence $\{\|x_i\|\}_{i \in \{k,\ldots,\infty\}}$ of positive numbers that $\sum_{i=k}^{\infty} \|x_i\|$ converges. Using 12.372 it follows then that $\sum_{i=k}^{\infty} \|x_{\beta(i)}\|$ is convergent. □

Next we must prove that the sum is independent of the bijection.

Theorem 12.389. *Let $\langle X, \|\|\rangle$ be a Banach space and $\sum_{i=k}^{\infty} x_i$ a absolute convergent serie, $\alpha, \beta: \{k, \ldots, \infty\} \rightarrow \{k, \ldots, \infty\}$ two bijections then $\sum_{i=k}^{\infty} x_{\alpha(i)} = \sum_{i=k}^{\infty} x_{\beta(i)}$ (the previous theorem ensures convergence). Note that by taking $\alpha = 1_{\{k,\ldots,\infty\}}$ we have that $\sum_{i=k}^{\infty} x_i = \sum_{i=k}^{\infty} x_{\beta(i)}$.*

Proof. Let $s = \sum_{i=k}^{\infty} x_{\alpha(i)}$ and $t = \sum_{i=k}^{\infty} x_{\beta(i)}$ and assume that $s \neq t \Rightarrow 0 < \|s - t\|$,. Then using 12.386 there exists finite $K_\alpha, K_\beta \subseteq \{k, \ldots, \infty\}$ such that $\forall H_\alpha, H_\beta$ finite with $K_\alpha \subseteq H_\alpha$ and $K_\beta \subseteq H_\beta$ we have

$$\left\| \sum_{i \in H_\alpha} x_{\alpha(i)} - s \right\|, \left\| \sum_{i \in H_\beta} x_i - t \right\| < \frac{\|s - t\|}{4} \tag{12.60}$$

Define now $P_{\alpha\beta} = K_\alpha \bigcup \alpha^{-1}(\beta(K_\beta))$ and $P_{\beta\alpha} = K_\beta \bigcup \beta^{-1}(\alpha(K_\alpha))$ then as α, β are bijections and K_α, K_β are finite we have that $P_{\alpha\beta}, P_{\beta\alpha}$ are finite and $H_\alpha \subseteq P_{\alpha\beta}$, $H_\beta \subseteq P_{\beta\alpha}$ so that by 12.60

$$\left\| \sum_{i \in P_{\alpha\beta}} x_{\alpha(i)} - s \right\|, \left\| \sum_{i \in P_{\beta\alpha}} x_{\beta(i)} - s \right\| < \frac{\|s - t\|}{4} \tag{12.61}$$

Now

$$\begin{aligned}\sum_{i\in P_{\alpha\beta}} x_{\alpha(i)} &= \sum_{i\in K_\alpha\bigcup\alpha^{-1}(\beta(K_\beta))} x_{\alpha(i)} \\ &\underset{10.44}{=} \sum_{i\in\alpha(K_\alpha\bigcup\alpha^{-1}(\beta(K_\beta)))} x_i \\ &= \sum_{i\in\alpha(K_\alpha)\bigcup\beta(K_\beta)} x_i \\ &= \sum_{i\in\beta(K_\beta\bigcup\beta^{-1}(\alpha(K_\alpha)))} x_i \\ &\underset{10.44}{=} \sum_{i\in K_\beta\bigcup\beta^{-1}(\alpha(K_\alpha))} x_{\beta(i)} \\ &= \sum_{i\in P_{\alpha\beta}} x_{\alpha(i)}\end{aligned}$$

proving that

$$\sum_{i\in P_{\alpha\beta}} x_{\alpha(i)} = \sum_{i\in P_{\alpha\beta}} x_{\alpha(i)} \tag{12.62}$$

Now using the above we have

$$\begin{aligned}\|s-t\| &= \left\|s-\sum_{i\in P_{\alpha\beta}} x_{\alpha(i)}+\sum_{i\in P_{\alpha\beta}} x_{\alpha(i)}-t\right\| \\ &\leqslant \left\|\sum_{i\in P_{\alpha\beta}} x_{\alpha(i)}-s\right\|+\left\|\sum_{i\in P_{\alpha\beta}} x_{\alpha(i)}-t\right\| \\ &<_{12.61} \frac{\|s-t\|}{4}+\frac{\|s-t\|}{4} \\ &< \frac{\|s-t\|}{2} \\ &< \|s-t\|\end{aligned}$$

a contradiction. So we must have that $s=t$. □

The above theorem lets us to extend the definition of a absolute convergent serie to a absolute convergent summable serie.

Definition 12.390. (Absolute convergent serie of a countable family) *Let $\langle X,\|\|\rangle$ be a Banach space, and I a countable set, $\{x_i\}_{i\in I}\subseteq X$ a countable family then we define absolute convergence of $\sum_{i\in I} x_i$ as follows*

1. *If I is finite then $\sum_{i\in I} x_i$ converges and $\sum_{i\in I} x_i$ is as defined in 10.37). Using 10.40 we have then that $\sum_{i\in I} x_i=\sum_{i=0}^{n} x_{\beta(i)}$ where $\beta:\{0,\ldots,n\}\rightarrow I$ is a bijection.*
2. *If I is infinite then $\sum_{i\in I} x_i$ converges if $\sum_{i=0}^{\infty} x_{\beta(i)}$ is absolute convergent where $\beta:\mathbb{N}_0\rightarrow I$ is a bijection (note that as I is inifinte and countable such a bijection exists). We denote then $\sum_{i\in I} x_i$ by $\sum_{i=0}^{\infty} x_{\beta(i)}$ (which is well defined as absolute convergence implies convergence (see 12.384))*

we must of course prove that this definition is independent of the choice of β

Proof.

1. This follows from 10.40
2. Let $\alpha: \mathbb{N}_0 \to I$ be a another bijection then $\alpha^{-1} \circ \beta: \mathbb{N}_0 \to \mathbb{N}_0$ is a bijection. So using 12.389 and absolute convergence we have that $\sum_{i=0}^{\infty} x_{\alpha(\alpha^{-1}(\beta(i)))}$ is convergent and $\sum_{i=0}^{\infty} x_{\alpha(i)} = \sum_{i=0}^{\infty} x_{\alpha(\alpha^{-1}(\beta(i)))} = \sum_{i=0}^{\infty} x_{\beta(i)} = s$ $\square$

Using the above definition we can extend 10.44

Theorem 12.391. *Let* $\langle X, \|\|\rangle$ *be a normed space* I *a countable set,* $\beta: I \to J$ *a bijection (so that* J *is also countable) and* $\sum_{j\in J} x_j$ *a absolute convergent space then* $\sum_{i\in I} x_{\beta(i)}$ *is absolute convergent and* $\sum_{i\in J} x_j = \sum_{i\in I} x_{\beta(i)}$

Proof. As I is countable we must consider two cases

***I* is finite.** then $\sum_{i\in I} x_{\beta(i)}$ is absolute convergent by definition.and $\sum_{i\in I} x_{\beta(i)} \underset{10.44}{=} \sum_{i\in J} x_j$

***I* is infinite.** then there exists a bijection $\alpha: \mathbb{N}_0 \to I$ so that $\beta \circ \alpha: \mathbb{N}_0 \to J$ is a bijection. By definition we have then that $\sum_{i=0}^{\infty} x_{\beta(\alpha(i))} = \sum_{i=0}^{\infty} x_{(\beta\circ\alpha)(i)}$ is absolute convergent. So we have that $\sum_{i\in I} x_{\beta(i)}$ is absolute convergent and $\sum_{i\in I} x_{\beta(i)} \underset{\text{definition}}{=} \sum_{i=1}^{\infty} x_{\beta(\alpha(i))} = \sum_{i=1}^{\infty} x_{(\beta\circ\alpha)(i)} \underset{\text{definition}}{=} \sum_{j\in J} x_j$ $\square$

Lemma 12.392. *Let* $\langle X, \|\|\rangle$ *be a normed space and* I, J *set such that there exists a bijection* $\beta: I \to J$ *then* $\{\sum_{i\in H} \|x_{\beta(i)}\| | H \subseteq J \wedge H \text{ is finite}\} = \{\sum_{i\in H} \|x_i\| | H \subseteq I \wedge H \text{ is finite}\}$

Proof.

$$s \in \left\{\sum_{i\in H} \|x_{\beta(i)}\| | H \subseteq I \wedge H \text{ is finite}\right\} \Leftrightarrow \exists H \text{ finite} \wedge H \subseteq I \text{ we have } s = \sum_{i\in H} \|x_{\beta(i)}\|$$
$$\Leftrightarrow \exists H \text{ finite} \wedge H \subseteq I \text{ we have } s \underset{10.44}{=} \sum_{i\in\beta(H)} \|x_i\|$$
$$\underset{\beta \text{ is bijective}}{\Leftrightarrow} \exists H \text{ finite} \wedge H \subseteq J \text{ we have } s = \sum_{i\in H} \|x_i\|$$
$$\Leftrightarrow s \in \left\{\sum_{i\in H} \|x_i\| | H \subseteq J \wedge H \text{ is finite}\right\}$$

$\square$

Theorem 12.393. *Let $\langle X, \|\|\rangle$ be a Banach space, I a countable set then $\sum_{i \in I} x_i$ is absolute convergent if and only if* $\sup(\{\sum_{i \in H} \|x_i\| | H \subseteq I \wedge H$ *is finite*$\})$ *exists.*

Proof.

⇒. as I is countable we must consider two cases

I is finite. then $\forall x \in \{\sum_{i \in H} \|x_i\| | H \subseteq I \wedge H \text{ is fiinite}\}$ we have that $\exists H$ finite with $H \subseteq I$ so that $x = \sum_{i \in H} \|x_i\| \leqslant \sum_{i \in I} \|x_i\| \in \{\sum_{i \in H} \|x_i\| | H \subseteq I \wedge H \text{ is fiinite}\}$ so that $\sup(\{\sum_{i \in H} \|x_i\| | H \subseteq I \wedge H \text{ is fiinite}\}) = \sum_{i \in I} \|x_i\|$ exists

I is infinite. then there exists a bijection $\beta: \mathbb{N}_0 \rightarrow I$ so that $\sum_{i=0}^{\infty} x_i$ is absolute convergent meaning that $\sum_{i=0}^{\infty} \|x_{\beta(i)}\|$ is convergent which by 12.375 gives that $\sup(\{\sum_{i \in H} \|x_{\beta(i)}\| | H \subseteq I \wedge H \text{ is finite}\})$ exists. Using 12.392 we have then that $\sup(\{\sum_{i \in H} \|x_i\| | H \subseteq I \wedge H \text{ is finite}))$ exists.

⇐. for I countable we have either

I is finite. then by definition $\sum_{i \in I} x_i$ is absolute convergent

I is infinite. then as I is countable there exists a bijection $\beta: \mathbb{N}_0 \rightarrow I$. By the hypothese $\sup(\{\sum_{i \in H} \|x_i\| | H \subseteq I \wedge H \text{ is finite}))$ exists so by 12.392 $\sup(\{\sum_{i \in \mathbb{N}_0} \|x_{\beta(i)}\| |\})$ exists which by 12.375 means the $\sum_{i=0}^{\infty} \|x_i\|$ is convergent or that $\sum_{i \in I} x_i$ is absolute convergent. □

We can now extend theorem 12.385 to this new definition of absolute convergence.

Theorem 12.394. *Let $\langle X, \|\|\rangle$ be a Banach space, I a countable set then $\sum_{i \in I} x_i$ is absolute convergent if and only if $\forall \varepsilon > 0$ there exists a finite $H \subseteq I$ such that $\forall K \subseteq I$ with $K \bigcap H = \emptyset$ we have $\sum_{i \in K} \|x_i\| < \varepsilon$*

Proof.

⇒. Take $\varepsilon > 0$, as I is countable we must consider two cases

I is finite. take $H = I$ then if $K \subseteq I$ with $K \bigcap H = \emptyset$ then $K = K \bigcap I = K \bigcap H = \emptyset$ so that $\sum_{i \in H} \|x_i\| = 0 < \varepsilon$

I is infinite. as $\sum_{i \in I} x_i$ is absolute convergent there exists a bijection $\beta: \mathbb{N}_0 \rightarrow I$ such that $\sum_{i=0}^{\infty} \|x_{\beta(i)}\|$ is absolute convergent. Using 12.385 there exists a finite $H' \subseteq \mathbb{N}_0$ such that $\forall K \subseteq \mathbb{N}_0$ with $K \bigcap H' = \emptyset$ we have $\sum_{i \in K} \|x_{\beta(i)}\| < \varepsilon$. Take $H = \beta(H')$ then H is finite and if $K \subseteq I$ with $K \bigcap H = \emptyset$ we have $\emptyset = \beta^{-1}(K \bigcap H) = \beta^{-1}(K) \bigcap \beta^{-1}(H) = \beta^{-1}(K) \bigcap H'$ so that $\sum_{i \in K} \|x_i\| \underset{10.44}{=} \sum_{i \in \beta^{-1}(K)} \|x_{\beta(i)}\| < \varepsilon$.

$\Leftarrow$. as I is countable we must consider two cases

I is finite. by definition we have then that $\sum_{i \in I} x_i$ is absolute convergent

I is infinite. then there exists a bijection $\beta: \mathbb{N}_0 \rightarrow I$. Let $\varepsilon > 0$, by the hypothese there exist a finite $H' \subseteq I$ such that $\forall K \subseteq H$ with $K \bigcap H = \emptyset$ we have $\sum_{i \in K} \|x_i\| < \varepsilon$. Take $H = \beta^{-1}(H')$ then if $K \subseteq \mathbb{N}_0$ with $K \bigcap H = \emptyset$ we have that $\emptyset = \beta(K \bigcap H) = \beta(K) \bigcap \beta(H) = \beta(K) \bigcap H'$ so that $\sum_{i \in K} \|x_{\beta(i)}\| \underset{10.44}{=} \sum_{i \in \beta(K)} \|x_i\| < \varepsilon$. □

We have now the following extension of the dominant convergence theorem.

Lemma 12.395. (Dominant Convergence) *Let $\langle X, \|\|\rangle$ be a Banach space, I a countable set, $\{x_i\}_{i \in I}$ a countable family such that there exists a $\{s_i\}_{i \in I} \subseteq [0, \ldots, \infty[$ such that $\sum_{i \in I} s_i$ is convergent and $\forall i \in I$ we have $\|x_i\| \leqslant s_i$ then $\sum_{i \in I} x_i$ is absolute convergent and $\|\sum_{i \in I} x_i\| \leqslant \sum_{i \in I} s_i$*

Proof. As I is countable we must consider two possibilities

I is finite. then by definition $\sum_{i \in I} x_i$ converges and $\|\sum_{i \in I} x_i\| \leqslant_{12.70} \sum_{i \in I} \|x_i\| \leqslant \sum_{i \in I} s_i$

I is infinite. then by definition (see 12.373) there exists a bijection $\beta: \mathbb{N}_0 \rightarrow I$ such that $\sum_{i=0}^{\infty} s_{\beta(i)}$ is convergent. As $\forall i \in \mathbb{N}_0$ we have that $|\|x_{\beta(i)}\|| = \|x_{\beta(i)}\| \leqslant s_{\beta(i)}$ we have using 12.381 that $\sum_{i=0}^{\infty} \|x_i\|$ is convergent and also that $\sum_{i=0}^{\infty} x_i$ is convergent with $\|\sum_{i=0}^{\infty} x_{\beta(i)}\| \leqslant \sum_{i=0}^{\infty} s_i$. So by definition $\sum_{i \in I} x_i$ is absolute convergent and $\|\sum_{i \in I} x_i\| = \|\sum_{i=0}^{\infty} x_{\beta(i)}\| \leqslant \sum_{i=0}^{\infty} s_{\beta(i)} = \sum_{i \in I} x_i$. □

We can extend now 12.386 to absolute convergent series of families

Lemma 12.396. *Let $\langle X, \|\|\rangle$ be a Banach space, I a countable set then $\sum_{i \in I} x_i$ is absolute convergent then $\forall \varepsilon > 0$ there exists a finite $K \subseteq I$ such that*

1. *$\forall H$ finite with $K \subseteq H \subseteq I$ we have that $\|\sum_{i \in I} x_i - \sum_{i \in H} x_i\| < \varepsilon$*
2. *$\forall L \subseteq I$ with L finite and $K \bigcap L = \emptyset$ we have $\sum_{i \in L} \|x_i\| < \varepsilon$*

Proof. *As I is countable we have two cases to consider*

I is finite. *take then $K = I$ so if $K \subseteq H \subseteq L$ then $H = I$ hence $\|\sum_{i \in I} x_i - \sum_{i \in I} x_i = \|\sum_{i \in I} x_i - \sum_{i \in H} x_i\| = 0 < \varepsilon$, further if $L \subseteq I$ with $K \bigcap L = \emptyset$ then $L = K \bigcap L = \emptyset$ so that $\sum_{i \in L} \|x_i\| = 0 < \varepsilon$*

I is infinite. *then by definition there exists a $\beta: \mathbb{N}_0 \rightarrow I$ such that $\sum_{i=0}^{\infty} x_{\beta(i)}$ is absolute convergent. Using 12.386 there exists then a $K' \subseteq \mathbb{N}_0$ such $\forall \varepsilon > 0$ we have*

1. *$\forall H$ finite with $K' \subseteq H \subseteq \mathbb{N}_0$ we have $\|\sum_{i=0}^{\infty} x_{\beta(i)} - \sum_{i \in H} x_{\beta(i)}\| < \varepsilon$*

2. $\forall L \subseteq \mathbb{N}_0$ *with* L *finite and* $K' \bigcap L = \emptyset$ *we have* $\sum_{i \in L} \|x_i\| < \varepsilon$

Define now $K = \beta(K')$ *then* $K \subseteq I$ *and* K *is finite. If* H *is a finite set such that* $K \subseteq H \subseteq I \underset{2.54}{\Rightarrow} \beta^{-1}(K) \subseteq \beta^{-1}(H) \subseteq \beta^{-1}(I) \Rightarrow K' \subset \beta^{-1}(H) \subseteq \mathbb{N}_0$ *hence using (1) we have* $\left\|\sum_{i=0}^{\infty} x_{\beta(i)} - \sum_{i \in \beta^{-1}(H)} x_{\beta(i)}\right\| < \varepsilon$. *As* $\sum_{i=0}^{\infty} x_{\beta(i)} \underset{\text{definition}}{=} \sum_{i \in I} x_i$ *and* $\sum_{i \in \beta^{-1}(H)} x_{\beta(i)} \underset{10.44}{=} \sum_{i \in B} x_i$ *we must have*

$$\forall H \text{ finite with } K \subseteq H \subseteq I \text{ we have } \left\|\sum_{i \in I} x_i - \sum_{i \in H} x_i\right\| < \varepsilon$$

Further if L *is finite with* $L \subseteq I$ *and* $K \bigcap L = \emptyset$ *then we have* $\beta^{-1}(L) \subseteq \beta^{-1}(I) = \mathbb{N}_0$ *and* $\beta^{-1}(K) \bigcap \beta^{-1}(L) = \beta^{-1}(\emptyset) = \emptyset \Rightarrow K' \bigcap \beta^{-1}(L) = \emptyset$ *so by (2) we have* $\sum_{i \in \beta^{-1}(L)} \|x_{\beta(i)}\| < \varepsilon$. *As* $\sum_{i \in \beta^{-1}(L)} \|x_{\beta(i)}\| \underset{10.44}{=} \sum_{i \in L} \|x_i\|$ *we conclude that*

$$\forall L \subseteq I \text{ with } L \text{ finete and } L \bigcap K = \emptyset \text{ we have } \sum_{i \in L} \|x_i\| < \varepsilon \qquad \square$$

We have the usually properties absolute convergent series of a countable family for scalar product and sum.

Theorem 12.397. *Let* $\langle X, \|\|\rangle$ *be a Banach space over a field* $\mathbb{K}$ *and* I *a countable set then we have*

1. *If* $\alpha \in \mathbb{K}$, $\sum_{i \in I} x_i$ *a absolute convergent serie then* $\sum_{i \in I} (\alpha \cdot x_i)$ *is absolute convergent and* $\sum_{i \in I} (\alpha \cdot x_i) = \alpha \cdot \sum_{i \in I} x_i$
2. *If* $\sum_{i \in I} x_i$, $\sum_{i \in I} y_i$ *are absoluite convergent series then* $\sum_{i \in I} (x_i + y_i)$ *is absolute convergent and* $\sum_{i \in I} (x_i + y_i) = \sum_{i \in I} x_i + \sum_{i \in I} y_i$

Proof. As I is countable we have either

***I* is finite.** then (1) follows from 10.123 and (2) follows from 10.53

***I* is infinite.**

1. then there exists as bijection $\beta: \mathbb{N}_0 \rightarrow I$ such that $\sum_{i=0}^{\infty} \|x_{\beta(i)}\|$ is convergent. Hence by 12.369 we have that $\sum_{i=0}^{\infty} \|\alpha \cdot x_i\| = \sum_{i=0}^{\infty} |\alpha| \cdot \|x_{\beta(i)}\|$ is convergent, proving that $\sum_{i \in I} (\alpha \cdot x_i)$ is absolute convergent (and thus convergent). Further we have $\sum_{i \in I} (\alpha \cdot x_i) \underset{\text{definition}}{=} \sum_{i=0}^{\infty} (\alpha \cdot x_{\beta(i)}) \underset{12.369}{=} \alpha \cdot \sum_{i=0}^{\infty} x_{\beta(i)} \underset{\text{definition}}{=} \alpha \cdot \sum_{i \in I} x_i$
2. then there exists as bijection $\beta: \mathbb{N}_0 \rightarrow I$ such that $\sum_{i=0}^{\infty} \|x_{\beta(i)}\|$, $\sum_{i=0}^{\infty} \|y_{\beta(i)}\|$ are convergent. Hence by 12.369 we have that $\sum_{i=0}^{\infty} \|x_{\beta(i)} + y_{\beta(i)}\|$ is convergent, proving that $\sum_{i \in I} (x_i + y_i)$ is absolute convergent (and thus convergent). Further we have $\sum_{i \in I} (x_i + y_i) \underset{\text{definition}}{=} \sum_{i=0}^{\infty} (x_{\beta(i)} + y_{\beta(i)}) \underset{12.369}{=} \sum_{i=0}^{\infty} x_{\beta(i)} + \sum_{i=0}^{\infty} y_{\beta(i)} \underset{\text{definition}}{=} \sum_{i \in I} x_i + \sum_{i \in I} y_i$ $\square$

Next we set out to prove associativity of absolute convergent series

Theorem 12.398. *Let $\langle X, \|\|\rangle$ be a Banach space, I, K countable sets, $\{N_i\}_{i\in I}$ a family of sets such that $\forall i, j \in I \vdash \bigcup_{i\in I} N_i = K$, $\sum_{i\in K} x_i$ a absolute convergent serie then $\forall i \in I$ we have that $\sum_{j\in N_i} x_j$ is absolute convergent and further $\sum_{i\in I} (\sum_{j\in N_i} x_j)$ is absolute convergent and $\sum_{i\in K} x_i = \sum_{i\in I} (\sum_{j\in N_i} x_j)$*

Proof. As $\sum_{i\in K} x_i$ is absolute convergent we have that $\sum_{i\in K} \|x_i\|$ is a convergent serie of positive numbers. So we can use the associativity of convergent series of positieve numbers (see 12.380) from which it follows that $\forall i \in I \vDash \sum_{j\in N_i} \|x_i\|$ converges and $\sum_{i\in I} (\sum_{j\in N_i} \|x_i\|)$ converges. Using dominant convergence (see 12.395) we have that

$$\forall i \in I \text{ that } \sum_{j\in N_i} x_j \text{ is absolute convergent and } \left\|\sum_{j\in N_i} x_j\right\| \leqslant \sum_{j\in N_i} \|x_j\| \tag{12.63}$$

Applying then dominant convergence on the last equality of the above we have that

$$\sum_{i\in I} \left(\sum_{j\in N_i} x_j\right) \text{ is absolute convergent} \tag{12.64}$$

Take now $s = \sum_{i\in K} x$ and $t = \sum_{i\in I} (\sum_{j\in N_i} z_j)$. Let $\varepsilon > 0$ then using the previous lemma (see 12.396) there exists a finite $G \subseteq K$ such that

$$\forall H \text{ finite with } G \subseteq H \subseteq K \text{ we have } \left\|\sum_{i\in H} x_i - s\right\| < \frac{\varepsilon}{3} \tag{12.65}$$

Define now $J = \{i \in I | N_i \bigcap G \neq \emptyset\} \subseteq I$. If $x \in G \subseteq K = \bigcup_{i\in I} N_i$ there exists a $i \in I$ such that $x \in G \bigcap N_i \Rightarrow i \in J \Rightarrow x \in \bigcup_{j\in J} (N_i \bigcap G)$ proving that $G \subseteq \bigcup_{j\in J} (N_i \bigcap G) \subseteq G$ so that $G = \bigcup_{j\in J} (N_i \bigcap G)$. As $J \subseteq K$ a countable set we have by 5.66 that J is countable. If now J is infinite we have as $\{N_i \bigcap G\}_{i\in J}$ is a pairwise disjoint countable family of non empty sets by 5.72 we would have that G is infinite a contradiction. Hence J is finite and thus $\bigcup_{i\in J} (N_i \bigcap G)$ is finite (as a finite union of finite sets (see 5.39)).

$$J = \left\{i \in I | N_i \bigcap G \neq \emptyset\right\} \text{ is finite and } G = \bigcup_{i\in J} \left(N_i \bigcap G\right) \tag{12.66}$$

As $\sum_{i\in I} (\sum_{j\in N_i} x_j) = s$ we can apply the previous lemma (see 12.396) again to find a finite $L \subseteq I$ such that $\forall H$ finite with $L \subseteq H \subseteq I$ we have $\|\sum_{i\in H} (\sum_{j\in N_i} x_j) - t\| < \frac{\varepsilon}{3}$. Take then $M = J \bigcup L$ then M is finite and $L \subseteq M \subseteq I$ so that

$$\left\|\sum_{i\in M} \left(\sum_{j\in N_i} x_j\right) - t\right\| < \frac{\varepsilon}{3} \tag{12.67}$$

As $\forall i \in M$ we have that $\sum_{j\in N_i} x_j$ is absolute convergent we have by using the previous lemma (see 12.396) that there exists a finite $M_i \subseteq N_i$ such that

$$\forall H \text{ finite with } M_i \subseteq H \subseteq N_i \text{ we have } \left\|\sum_{j\in H} x_j - \sum_{j\in N_i} x_j\right\| < \frac{\varepsilon}{3 \cdot \#(M)} \tag{12.68}$$

Now $\forall i\in M$ define $K_i = M_i \bigcup (N_i \bigcap G)$ a finite set then we have that $M_i \subseteq K_i \subseteq N_i$ so by 12.68 we have

$$\forall i\in M \text{ we have } \left\| \sum_{j\in K_i} x_j - \sum_{j\in N_i} x_j \right\| < \frac{\varepsilon}{3\cdot\#(M)} \tag{12.69}$$

Now as $G \underset{12.66}{=} \bigcup_{j\in J} (G\bigcap N_j) \subseteq \bigcup_{j\in M} (G\bigcap N_j) \subseteq \bigcup_{j\in M} (J_j \bigcup (G\bigcap N_j)) = \bigcup_{j\in M} K_j \subseteq_{J_j\subseteq N_j} N_i$ and as M is finite we have that $\bigcup_{j\in M} K_j$ is finite. So using 12.65 we have that

$$\left\| \sum_{i\in\bigcup_{j\in M} K_j} x_i - s \right\| < \frac{\varepsilon}{3} \tag{12.70}$$

Let's calculate now $\|s-t\|$

$$\begin{aligned}
\|s-t\| &= \left\| s - \sum_{i\in\bigcup_{j\in M}K_j} x_i - \sum_{i\in\bigcup_{j\in M}K_j} x_i - t \right\| \\
&\leqslant \left\| s - \sum_{i\in\bigcup_{j\in M}K_j} x_i \right\| + \left\| \sum_{i\in\bigcup_{j\in M}K_j} x_i - t \right\| \\
&<_{12.70} \frac{\varepsilon}{3} + \left\| \sum_{i\in\bigcup_{j\in M}K_j} x_i - t \right\| \\
&\underset{10.47+\text{if } i\neq j \Rightarrow K_i\bigcap K_j=\emptyset}{=} \frac{\varepsilon}{3} + \left\| \sum_{j\in M} \left(\sum_{j\in K_j} x_j \right) - t \right\| \\
&= \frac{\varepsilon}{3} + \left\| \sum_{i\in M} \left(\sum_{j\in K_i} x_j \right) - \sum_{i\in M} \left(\sum_{j\in N_i} x_j \right) + \right. \\
&\quad \left. \sum_{i\in M} \left(\sum_{j\in N_i} x_j \right) - t \right\| \\
&\leqslant \frac{\varepsilon}{3} + \left\| \sum_{i\in M} \left(\sum_{j\in K_i} x_j \right) - \sum_{i\in M} \left(\sum_{j\in N_i} x_j \right) \right\| + \\
&\quad \left\| \sum_{i\in M} \left(\sum_{j\in N_i} x_j \right) - t \right\| \\
&<_{12.67} \frac{\varepsilon}{3} + \left\| \sum_{i\in M} \left(\sum_{j\in K_i} x_j \right) - \sum_{i\in M} \left(\sum_{j\in N_i} x_j \right) \right\| + \frac{\varepsilon}{3} \\
&= \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \left\| \sum_{i\in M} \left(\sum_{j\in K_i} x_j - \sum_{j\in N_i} x_j \right) \right\| \\
&\leqslant_{M \text{ is finite}} \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \sum_{i\in M} \left\| \sum_{j\in K_i} x_j - \sum_{j\in N_i} x_j \right\|
\end{aligned}$$

$$\begin{array}{ll} \leqslant_{12.69} & \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \sum_{i \in M} \left(\frac{\varepsilon}{3 \cdot \#(M)} \right) \\ = & \varepsilon \end{array}$$

So given $\varepsilon > 0$ we have that $\|s - t\| < \varepsilon$ from which it follows that $\|s - t\| \leqslant 0$ and thus $s = t$ or using the definitions of s and t that

$$\sum_{i \in K} x_i = \sum_{i \in I} \left(\sum_{j \in N_i} x \right) j$$

□

As a application of associativity we proof now distributivity for absolute convergent series in $\mathbb{R}$

Theorem 12.399. *Let I, J be countable sets, $\langle \mathbb{R}, || \rangle$ the set of the reals with the absolute value as norm (which is a Banach space by 12.348). $\sum_{i \in I} x_i$ and $\sum_{i \in J} y_i$ be absolute convergent series of countable families then $\sum_{(i,j) \in I \times J} (x_i \cdot y_j)$ is absolute convergent and $\sum_{(i,j) \in I \times J} (x_i \cdot y_j) = (\sum_{i \in I} x_i) \cdot (\sum_{j \in J} y_j)$ [note that by 5.69 $I \times J$ is countable]. Here $\sum_{(i,j) \in I \times J} (x_i \cdot y_j)$ is a shorthand for $\sum_{\alpha \in I \times J} (x_{\pi_1(\alpha)}, y_{\pi_2(\alpha)})$*

Proof. As $\sum_{i \in I} x_i$ and $\sum_{j \in J} y_j$ are are absolute convergent we have that $s = \sup\left(\{\sum_{i \in I} \|x_i\| | H \subseteq I \wedge H \text{ finite}\}\right)$ and $t = \sup\left(\{\sum_{j \in H} \|y_j\| | H \subseteq J \wedge H \text{ finite}\}\right)$ exists (see 12.393). Take now a finite $H \subseteq I \times J$ then as $\pi_1: H \rightarrow \pi_1(H)$ and $\pi_2: H \rightarrow \pi_2(H)$ are surjective we have that $\pi_1(H), \pi_2(H)$ are finite (see 5.48). If $(i, j) \in H$ then $i = \pi_1(i) \in \pi_1(H)$ and $j = \pi_2(i, j) \in \pi_2(H)$ so that $(i, j) \in \pi_1(H) \times \pi_2(H)$ proving that $H \subseteq \pi_1(H) \times \pi_2(H)$. Hence $\sum_{(i,j) \in H} (\|x_i\| \cdot \|y_j\|) \leqslant \sum_{(i,j) \in \pi_1(I) \times \pi_2(J)} =_{10.60}$

$$\begin{array}{rcl} \sum_{(i,j) \in H} \|x_i \cdot y_j\| & = & \sum_{(i,j) \in H} (\|x_i\| \cdot \|y_j\|) \\ & \leqslant & \sum_{(i,j) \in \pi_1(H) \times \pi_2(H)} (\|x_i\| \cdot \|y_j\|) \\ & =_{10.60} & \left(\sum_{i \in \pi_1(H)} \|x_i\| \right) \cdot \left(\sum_{j \in p_2(H)} \|y_j\| \right) \\ & \leqslant_{9.41} & s \cdot t \end{array}$$

So $\sup\left(\left\{\sum_{(i,j) \in H} \|x_i \cdot y_j\| | H \subseteq I \times J \wedge H \text{ finite}\right\}\right)$ exists (see 9.43) and thus by 12.393 we have that

$$\sum_{(i,j) \in I \times J} (x_i \cdot y_j) \text{ is absolute convergent} \tag{12.71}$$

Next we prove that

$$I \times J = \bigcup_{i \in I} (\{i\} \times J) \text{ and } \forall i, j \in I \vdash \text{ we have } \{i\} \times J \bigcap \{j\} \times J = \emptyset \tag{12.72}$$

First if $(x,y)\in\bigcup_{i\in I}(\{i\}\times J)$ then $\exists i\in I$ such that $(x,y)\in\{i\}\times J\Rightarrow x=i\in I\wedge y\in J=(x,y)\in I\times J$. Second if $(x,y)\in I\times J$ then $x\in I\wedge j\in J\Rightarrow(x,y)\in\{x\}\times J\wedge x\in I\Rightarrow(x,y)\in\bigcup_{i\in I}\{i\}\times J$ proving $I\times J=\bigcup_{i\in I}(\{i\}\times J)$. Also if $i\neq j$ and $(x,y)\in\{i\}\times J\bigcap\{j\}\times J$ then $i=x=j$ contradicting $i\neq j$. this proves our assertion.

Let $i\in I$ and take $(x,y),(x',y')\in\{i\}\times J$ such that $\pi_2((x,y))=\pi_2((x',y'))$ then $x=i=x'$ and $y=y'$ so that $(x,y),(x',y')$. Further if $j\in J$ then $(i,j)\in\{i\}\times J$ and $\pi_2((i,j))=j$. So we have that

$$\forall i\in I\ \pi_2:\{i\}\times J\rightarrow J \text{ is a bijection} \tag{12.73}$$

$$\begin{aligned}
\sum_{\alpha\in I\times J}(x_{\pi_1(\alpha)}\cdot y_{\pi_2(\alpha)}) &\underset{12.72 \text{ and asociativity (see12.398)}}{=} \sum_{i\in I}\left(\sum_{\alpha\in\{i\}\times J}(x_{\pi_1(\alpha)}\cdot y_{\pi_2(\alpha)})\right)\\
&\underset{a\in\{i\}\times J\Rightarrow\pi_1(\alpha)=i}{=} \sum_{i\in I}\left(\sum_{\alpha\in\{i\}\times J}\right)(x_i\cdot y_{\pi_2(\alpha)})\\
&\underset{12.397}{=} \sum_{i\in I}\left(x_i\cdot\sum_{\alpha\in\{i\}\times J}y_{\pi_2(\alpha)}\right)\\
&\underset{12.73 \text{ together with } 12.391}{=} \sum_{i\in I}\left(x_i\cdot\sum_{j\in J}y_j\right)\\
&\underset{12.397}{=} \left(\sum_{i\in I}x_i\right)\cdot\left(\sum_{j\in J}y_j\right)
\end{aligned}$$

□

12.11.8 Properties of complete spaces

Definition 12.400. *Let $\langle S,d\rangle$ be a metric space then a function $f:S\rightarrow S$ is a* ***contraction*** *if $\exists\lambda\in[0,1[\subseteq\mathbb{R}$ such that $\forall x,y\in S$ we have $d(f(x),f(y))\leqslant\lambda\cdot d(x,y)$*

Lemma 12.401. *Let $\langle S,d\rangle$ be a metric space and $f:S\rightarrow S$ a contraction then f is continuous*

Proof. Let $\lambda\in[0,1[$ be such that $d(f(x),f(y))\leqslant\lambda\cdot d(x,y)$ then given a $x\in S$ and a $y\in S$ such that $d(x,y)<\varepsilon$ we have $d(f(x),f(y))\leqslant\lambda\cdot d(x,y)<\lambda\cdot d(x,y)<d(x,y)$. □

Lemma 12.402. *Let $\lambda\in\mathbb{R}$ such that $\lambda\neq 1$ then if $k\in\mathbb{N}$ and $n\in\mathbb{N}_0$ we have $\sum_{i=1}^k\lambda^{n+(i-1)}=\lambda^n\cdot\frac{1-\lambda^k}{1-\lambda}$. So if we take $n=0$ then $\sum_{i=1}^k\lambda^{(i-1)}=\frac{1-\lambda^k}{1-\lambda}$ and if $n=1$ then $\sum_{i=1}^k\lambda^i=\lambda\cdot\frac{1-\lambda^k}{1-\lambda}$*

Proof. We proof this by induction so let $S=\left\{k\in\mathbb{N}|\sum_{i=1}^k\lambda^{n+(i-1)}=\lambda^n\cdot\frac{1-\lambda^k}{1-\lambda}\right\}$ then we have:

1. If $k=1$ then $\sum_{i=1}^1\lambda^{n+(i-1)}=\lambda^n=\lambda^n\cdot\frac{1-\lambda}{1-\lambda}=\lambda^n\cdot\frac{1-\lambda^1}{1-\lambda}$ so that $1\in S$

2. Assume that $k \in S$ then we have

$$\begin{aligned}\sum_{i=1}^{k+1} \lambda^{n+(i-1)} &= \left(\sum_{i=1}^{k} \lambda^{n+(i-1)}\right) + \lambda^{n+((k+1)-1)} \\ &= \left(\sum_{i=1}^{k} \lambda^{n+(i-1)}\right) + \lambda^{n+k} \\ &\underset{k\in S}{=} \lambda^n \cdot \frac{1-\lambda^k}{1-\lambda} + \lambda^{n+k} \\ &= \frac{\lambda^n \cdot (1-\lambda^k) + (1-\lambda) \cdot \lambda^{n+k}}{1-\lambda} \\ &= \frac{\lambda^n \cdot (1-\lambda^k + (1-\lambda) \cdot \lambda^k)}{1-\lambda} \\ &= \frac{\lambda^n \cdot (1-\lambda^k + \lambda^k - \lambda^{k+1})}{1-\lambda} \\ &= \lambda^n \cdot \frac{1-\lambda^{k+1}}{1-\lambda}\end{aligned}$$

proving that $k+1 \in S$ □

Lemma 12.403. *Let $\lambda \in \mathbb{R}$ with $0 < \lambda < 1$ then $\sum_{i=1}^{\infty} \lambda^i$ converges and $\sum_{i=1}^{\infty} \lambda^i = \frac{\lambda}{1-\lambda}$*

Proof. Given a $k \in \mathbb{N}$ we have that $\left|\left(\sum_{i=1}^{k} \lambda^i\right) - \frac{\lambda}{1-\lambda}\right| = \left|\left(\lambda \cdot \sum_{i=1}^{k} \lambda^{i-1}\right) - \frac{\lambda}{1-\lambda}\right| \underset{\text{previous lemma with } n=0}{=} \left|\lambda \cdot \left(\lambda^0 \cdot \frac{1-\lambda^k}{1-\lambda}\right) - \frac{\lambda}{1-\lambda}\right| = \left|\frac{\lambda}{1-\lambda} \cdot (1-\lambda^k - 1)\right| = \left|\frac{\lambda^{k+1}}{1-\lambda}\right| = \frac{\lambda}{1-\lambda} \cdot \lambda^k$. Now using 12.319 given $\varepsilon > 0$ there exists a $N \in \{1, ..., \infty\}$ such that $\frac{\lambda}{1-\lambda} \cdot \lambda^k = \left|\frac{\lambda}{1-\lambda} \cdot \lambda^k - 0\right| < \varepsilon$ if $k \geqslant N \Rightarrow \left|\left(\sum_{i=1}^{k} \lambda^i\right) - \frac{\lambda}{1-\lambda}\right| < \varepsilon$ if $k \geqslant N$ □

Theorem 12.404. (The Banach Fixed Point Theorem) *Let $\langle S, d\rangle$ be a complete metric space and $f: S \rightarrow S$ any* ***contraction*** *then $\exists! x_0 \in S$ such that $f(x_0) = x_0$ (f has exactly one fixed point x_0). If $x \in S$ then $\lim_{n\rightarrow\infty} f^{(n)}(x) \rightarrow x_0$ where f^n is recursively defined by*

$$\begin{aligned}f^{(1)} &= f \\ f^{(n+1)} &= f \circ f^{(n)}\end{aligned}$$

Proof. First as f is a contraction there exists a $\lambda \in [0, 1[$ such that $d(f(x), f(y)) \leqslant \lambda \cdot d(x, y)$. As S is not empty there exists a $x \in S$. We prove now by induction that $\forall n \in \mathbb{N}$ we have $d(f^{(n)}(x), f^{(n+1)}(x)) \leqslant \lambda^n \cdot d(x, f(x))$. So take then $P = \{n \in \mathbb{N} | d(f^{(n)}(x), f^{(n+1)}(x)) \leqslant \lambda^n \cdot d(x, f(x))\}$ then we have:

1. If $n = 1$ then $d(f^{(n)}(x), f^{(n+1)}(x)) = d(f^{(1)}(x), f^{(2)}(x)) = d(f(x), f((f(x)))) \leqslant \lambda \cdot d(x, f(x)) = \lambda^1 \cdot d(x, f(x))$ proving that $1 \in P$

2. If $n \in P$ then we have $d(f^{(n+1)}(x), f^{((n+1)+1)}(x)) = d(f(f^{(n)}(x)), f(f^{(n+1)}(x))) \leqslant \lambda \cdot d(f^{(n)}(x), f^{(n+1)}(x)) \leqslant \lambda \cdot \lambda^n \cdot d(x, f(x)) = \lambda^n \cdot d(x, f(x))$ proving that $n+1 \in P$

Next we prove by induction that given $n, k \in \mathbb{N}$ that $d(f^{(n)}(x), f^{(n+k)}(x)) \leqslant (\sum_{i=1}^{k} \lambda^{n+(i-1)}) \cdot d(x, f(x))$. So let $Q_n = \{k \in \mathbb{N} | d(f^{(n)}(x), f^{(n+k)}(x)) \leqslant (\sum_{i=1}^{k} \lambda^{n+(i-1)}) \cdot d(x, f(x))\}$ then we have:

1. If $k = 1$ then $d(f^{(n)}(x), f^{(n+1)}(x)) \leqslant \lambda^n \cdot d(x, f(x)) = (\sum_{i=1}^{1} \lambda^{n+(i-1)}) \cdot d(x, f(x))$ proving that $1 \in Q_n$

2. If $k \in Q_n$ then

$$\begin{aligned}
d(f^{(n)}(x), f^{(n+(k+1))}(x)) &\leqslant d(f^{(n)}(x), f^{(n+k)}(x)) + d(f^{(n+k)}(x), f^{(n+(k+1))}(x)) \\
&\leqslant \left(\sum_{i=1}^{k} \lambda^{n+(i-1)}\right) \cdot d(x, f(x)) + d(f^{(n+k)}(x), f^{((n+k)+1)}(x)) \\
&\leqslant \left(\sum_{i=1}^{k} \lambda^{n+(i-1)}\right) \cdot d(x, f(x)) + \lambda^{n+k} \cdot d(x, f(x)) \\
&= \left(\left(\sum_{i=1}^{k} \lambda^{n+(i-1)}\right) + \lambda^{n+((k+1)-1)}\right) \cdot d(x, f(x)) \\
&= \left(\sum_{i=1}^{k+1} \lambda^{n+(i-1)}\right) \cdot d(x, f(x))
\end{aligned}$$

proving that $k+1 \in Q_n$

So $d(f^{(n)}(x), f^{(n+k)}(x)) \leqslant (\sum_{i=1}^{k+1} \lambda^{n+(i-1)}) \cdot d(x, f(x)) \underset{12.402}{=} \lambda^n \cdot \frac{1-\lambda^k}{1-\lambda} \cdot d(x, f(x)) <_{\lambda^k > 0} \frac{\lambda^n}{1-\lambda} \cdot d(x, f(x))$ proving

$$d(f^{(n)}(x), f^{(n+k)}(x)) < \frac{\lambda^n}{1-\lambda} \cdot d(x, f(x)) \tag{12.74}$$

Next we prove that $\{f^{(n)}(x)\}_{n \in \{1, \ldots, \infty\}}$ is Cauchy (using 12.310) so let $\varepsilon > 0$ be given then using 12.74 and 12.319 there exists a $N \in \{1, \ldots, \infty\}$ such that if $k \geqslant N$ then $d(f^{(n)}(x), f^{(n+k)}(x)) < \varepsilon$ proving the Cauchy property.

As S is complete there exists a $x_0 \in S$ such that

$$\lim_{n \to \infty} f^{(n)}(x) = x_0 \tag{12.75}$$

Using the fact that a contraction mapping is continuous (12.401), 12.75 and 12.323 we have

$$\lim_{n \to \infty} f^{(n+1)}(x) = \lim_{n \to \infty} f(f^{(n)}(x)) = f(x_0) \tag{12.76}$$

Assume now that $f(x_0) \neq x_o$ then $d(f(x_0), x_o) > 0$ by 12.75 there exists a N_1 such that $d(f^{(n)}(x)) < \frac{d(f(x_0), x_0)}{2}$ if $n \geqslant N_1$, likewise from 12.76 there exists a N_2 such that $d(f^{(n+1)}(x), f(x)) < \frac{d(f(x_0), x_0)}{2}$. Let $n \geqslant \max(N_1, N_2 + 1) \Rightarrow n \geqslant N_1, n - 1 \geqslant N_2$ then $d(f(x_0), x_0) \leqslant d(f(x_0), f^{((n-1)+1)}(x)) + d(f^{((n-1)+1)}(x), x_0) = d(f(x_0), f^{((n-1)+1)}(x)) + d(f^{(n)}(x), x_0) < \frac{d(f(x_0), x_0)}{2} + \frac{d(f(x_0), x_o)}{2} = d(f(x_0), x_0)$ giving the contradiction $d(f(x_0), x_0) < d(f(x_0), x_0)$. So we must have that $f(x_0) = x_0$.

Finally to prove uniqueness assume that there is also a $x_1 \in S$ such that $f(x_1) = x_1$, then if $x_1 \neq x_0$ then $0 < d(x_1, x_0) = d(f(x_0), f(x_1)) < \lambda d(x_0, x_1) < d(x_0, x_1)$ as $\lambda \in [0, 1[\Rightarrow d(x_1, x_0) < d(x_1, x_0)$ a contradiction so $x_0 = x_1$ □

Lemma 12.405. *If $\{A_i\}_{i \in \mathbb{N}}$ is a family of sets such that $\forall i \in \mathbb{N} \setminus \{1\}$ we have that $A_i \subseteq A_{i-1}$ then for every $n, m \in \mathbb{N}$ with $n \leqslant m$ we have $A_m \subseteq A_n$*

Proof. This is easy proved by induction. So let $S_n = \{k \in \mathbb{N}_0 | A_{n+k} \subseteq A_n\}$ then we have

1. If $k = 0$ then $A_{n+k} = A_n \subseteq A_n$ proving that $0 \in S_n$
2. If $k \in S_n$ then we have that $A_{n+(k+1)} = A_{(n+k)+1} \subseteq A_{((n+k)+1)-1} = A_{n+k} \subseteq A_n \Rightarrow k + 1 \in S_n$

Using induction we have then that $S_n = \mathbb{N}_0$. If now $n \leqslant m$ then $m - n \in \mathbb{N}_0 = S_n$ so that $A_m = A_{n+(m-n)} \subseteq A_n$ □

Lemma 12.406. *Let $\langle X, d\rangle$ be a complete metric space and let $\{A_i\}_{i \in \mathbb{N}}$ be a sequence of non empty closed sets with $\forall i \in \mathbb{N} \setminus \{1\}$ we have $A_i \subseteq A_{i-1}$ and $\lim_{n \to \infty} \operatorname{diam}(A_i) = 0$ then $\bigcap_{i \in \mathbb{N}} A_i$ is not empty.*

Proof. Using the axiom of choice (see 2.201) there exists a function $c: \mathbb{N} \to \bigcup_{i \in \mathbb{N}} A_i$ such that $\forall i \in \mathbb{N}$ we have $c_i \in A_i$ defining the family $\{c(i)\}_{i \in \mathbb{N}}$. Now if $\varepsilon > 0$ we use $\lim_{n \to \infty} \operatorname{diam}(A_i) = 0$ to find a $N \in \mathbb{N}$ such that $\forall n \geqslant N$ we have $\operatorname{diam}(A_n) = |\operatorname{diam}(A_n) - 0| < \varepsilon$. If now $n, m \geqslant N$ then using the above lemma we have that A_n, $A_m \subseteq A_N$ so that $c(n), c(m) \in A_N$ and thus $d(c(n), c(m)) \leqslant \operatorname{diam}(A_n) < \varepsilon$ proving that $\{c(i)\}_{i \in \mathbb{N}}$ is Cauchy. As X is complete there exists a $x \in X$ such that $\lim_{i \to \infty} c(i) = x$. If now $n \in \mathbb{N}$ and U a open set with $x \in U$ then there exists a $\varepsilon > 0$ such that $x \in B_d(x, \varepsilon) \subseteq U$. From the definition of x there exists then a $N \in \mathbb{N}$ such that if $m \geqslant N$ we have $d(x, c(m)) < \varepsilon \Rightarrow c(m) \in B_d(x, \varepsilon) \subseteq U$. So if we take $q = \max(N, n)$ then $q \geqslant n \Rightarrow c(q) \in A_n$ and $q \geqslant N \Rightarrow c(q) \in U$ proving that $A_n \bigcap U \neq \emptyset$ so that $x \in A = \overline{A_n} \underset{A \text{ is closed}}{=} A$. As n was chosen arbitrary we have that $\forall n \in \mathbb{N}$ that $x \in A_n$ or $x \in \bigcap_{n \in \mathbb{N}} A_n$. □

Theorem 12.407. (Baire Category Theorem) *If X is a compact Hausdorff topological space or if $\langle X, d\rangle$ is a complete metric space then X is a Baire space (12.42).*

Proof. First note that because of 12.222 and 12.248 we have that X is regular in both cases. Let $\mathcal{T}$ be the topology of $\langle X, d\rangle$ and take then $\mathcal{A} = \{U \in \mathcal{T} | U \neq \emptyset\}$. Let $\{A_i\}_{i \in \mathbb{N}}$ be a sequence of closed sets of X with $\forall i \in \mathbb{N}$ we have $A_i^\circ = \emptyset$, we must show then that $\left(\bigcup_{i \in \mathbb{N}} A_i\right)^\circ = \emptyset$. Take $U_0 \in \mathcal{A}$ a non empty open set arbitrarily.

Consider the set A_1 then because $A_1^0 = \emptyset$ we have $U_0 \nsubseteq A_1 \Rightarrow \exists y \in U_0$ such that $y \notin A_1$, using regularity there exists open sets V_1, U_1' such that $y \in U_1' \wedge A_1 \subseteq V_1 \wedge U_1' \bigcap V_1 = \emptyset$. From $U_1' \bigcap V_1 = \emptyset$ we deduce that $U_1' \subseteq X \setminus V_1$ which is closed so that $\overline{U_1'} \subseteq X \setminus V_1$ and thus $\overline{U_1'} \bigcap V_1 = \emptyset \Rightarrow \overline{U_1'} \bigcap A \subseteq \overline{U_1'} \bigcap V_1 = \emptyset$. Using 12.223 there exists a U_1'' such that $y \in U_1''$ and $\overline{U''} \subseteq U_0$. If we take now $U_1 = U_1' \bigcap U'' \bigcap U_0 \ni y \Rightarrow U_1 \in \mathcal{A}$ [in case of a complete metric space we take $U_1 = U_1' \bigcap U'' \bigcap B_d\left(y, \frac{1}{2}\right) \ni y \Rightarrow U_1 \in \mathcal{A}$] giving

$$\overline{U_1} \bigcap A_1 = \emptyset \wedge \overline{U_1} \subseteq U_0 \text{ and in case of } a \text{ metric space } \operatorname{diam}(\overline{U_1}) \leqslant 1 \tag{12.77}$$

Define now $\mathcal{N}_1 = \left\{\{U\}_{i \in \{1, \ldots, n\}} | n \in \mathbb{N} \wedge \forall i \in \{1, \ldots, n\} \text{ we have } U_i \in \mathcal{A} \wedge \text{we have } \bar{U}_i \bigcap A_i = \emptyset \wedge \bar{U}_i \subseteq U_{i-1} \text{ and in case of } a \text{ metric space } \operatorname{diam}(\bar{U}_i) \leqslant \frac{1}{i}\right\} \subseteq \mathcal{M} = \{\{U_i\}_{i \in \{0, \ldots, n\}} | n \in \mathbb{N}_0 \wedge \forall i \in \{0, \ldots, n\} \text{ we have } U_i \in \mathcal{A}\}$. Using 12.77 we have then

$$\{U_1\}_{i \in \{1, \ldots, 1\}} \in \mathcal{N}_1 \tag{12.78}$$

Define now $\rho: \mathcal{N}_1 \rightarrow \mathcal{A}$ as follows. Take $\{U\}_{i \in \{1, \ldots, n\}} \in \mathcal{N}_1$ then as U_n is not empty and $A_{n+1}^\circ = \emptyset$ we can not have $U_n \subseteq A_{n+1} \Rightarrow \exists y \in U_n$ such that $y \in A_{n+1}$, using regularity we can find a $U_{n+1}', V_{n+1} \in \mathcal{T}$ such that $y \in U_{n+1}', A_{n+1} \subseteq V_{n+1}$ and $U_{n+1}' \bigcap V_{n+1} = \emptyset$. So $U_{n+1}' \subseteq X \setminus V_{n+1}$ a closed set so that $\overline{U_{n+1}'} \subseteq X \setminus V_{n+1}$ giving $\overline{U_{n+1}} \bigcap V_{n+1} = \emptyset$. Using 12.223 there exists a $U_{n+1}'' \in \mathcal{T}$ with $y \in U_{n+1}''$ and $\overline{U_{n+1}''} \subseteq U_{n+1}'$. Taken now $\rho(\{U_i\}_{i \in \{1, \ldots, n\}}) = U_{n+1}' \bigcap U_{n+1}'' \bigcap U_n \ni y$ [in case of a metric space we take $\rho(\{U_i\}_{i \in \{1, \ldots, n\}}) = U_{n+1}' \bigcap U_{n+1}'' \bigcap U_n \bigcap B_d\left(y, \frac{1}{2 \cdot (n+1)}\right) \ni y$ so that we have that

$$\rho(\{U_i\}_{i \in \{1, \ldots, n\}}) \in \mathcal{A} \wedge \overline{\rho(\{U_i\}_{i \in \{1, \ldots, n\}})} \bigcap A_{n+1} = \emptyset \wedge \overline{\rho(\{U_i\}_{i \in \{1, \ldots, n\}})} \subseteq U_n \text{ and in case of } a \text{ metric space } \operatorname{diam}(\overline{\rho(\{U_i\}_{i \in \{1, \ldots, n\}})}) \leqslant \frac{1}{n} \tag{12.79}$$

So if we define $\{U_i'\}_{i \in \{1, \ldots, n+1\}}$ by $U_i' = \begin{cases} U_i \text{ iof } i \in \{1, \ldots, n\} \\ \rho(\{U_i\}_{i \in \{1, \ldots, n\}}) \text{ if } i = n+1 \end{cases}$ then we have

$$\{U_i'\}_{i \in \{1, \ldots, n+1\}} \in \mathcal{N}_1 \tag{12.80}$$

Using 5.26 there exists a function $U: \mathbb{N} \rightarrow \mathcal{A}$ such that

1. $U(1) = U_1$
2. $\forall n \in \mathbb{N}$ we have $U(n+1) = \rho(\{U(i)\}_{i \in \{1, \ldots, n\}})$
3. $\forall n \in \mathbb{N}$ we have $\{U(i)\}_{i \in \{1, \ldots, n\}} \in \mathcal{N}_1$

This defines the sequence $\{U(i)\}_{i \in \mathbb{N}}$ in $\mathcal{A}$ (the set of non empty open sets) so that:

$$\forall i \in \mathbb{N} \text{ we have } \overline{U(i)} \bigcap A_i = \emptyset, \bar{U}_i \subseteq U_{i-1}, \text{ in case of } a \text{ metric space } \operatorname{diam}(\overline{U(i)}) \leqslant \frac{1}{i} \tag{12.81}$$

We prove now that $\bigcap_{i \in \mathbb{N}} \bar{U}_i \neq \emptyset$ consider the two cases:

1. **(X is compact Hausdorff)** Then we prove by induction that $\emptyset \neq \overline{U_n} \subseteq \bigcap_{i \in \{1,\ldots,n\}} \bar{U}_i$. So let $\mathcal{S} = \left\{n \in \mathbb{N} | \overline{U_n} \subseteq \bigcap_{i \in \{1,\ldots,n\}} \bar{U}_i\right\}$ then we have:

   a. If $n = 1$ then $\overline{U_1} \subseteq \overline{U_1} = \bigcap_{i \in \{1,\ldots,1\}} \bar{U}_i$ so that $1 \in \mathcal{S}$

   b. If $n \in \mathcal{S}$ then $\overline{U_{n+1}} \subseteq U_n \subseteq \overline{U_n} \Rightarrow \overline{U_{n+1}} \subseteq \overline{U_n} \bigcap \overline{U_{n+1}} \subseteq \left(\bigcap_{i \in \{1,\ldots,n\}} \bar{U}_i\right) \bigcap \overline{U_{n+1}} = \bigcap_{i \in \{1,\ldots,n+1\}} \bar{U}_i$ proving that $n + 1 \in S$.

   Mathematical induction proves then that $S = \mathbb{N}_{0\,1}$. This proves that $\{\bar{U}_i\}_{i \in \mathbb{N}}$ has the finite intersection property so that by 12.286 we have that $\bigcap_{i \in \mathbb{N}} \bar{U}_i \neq \emptyset$

2. **($\langle X, d\rangle$ is a metric space)** As $\mathrm{diam}(\bar{U}_i) \leqslant \frac{1}{i}$ we have that $\lim_{i \to \infty} \mathrm{diam}(\bar{U}_i) = 0$ so by the previous lemma and $\bar{U}_i \subseteq U_{i-1} \subseteq \overline{U_{i-1}}$ we have again that $\bigcap_{i \in \mathbb{N}} \bar{U}_i \neq \emptyset$.

Take now $x \in \bigcap_{i \in \mathbb{N}} \bar{U}_i$ then $x \in \overline{U_1} \subseteq U_0 \Rightarrow x \in U_0$ and $\forall i \in \mathbb{N}$ we have $\bar{U}_i \bigcap A_i = \emptyset$ so that $x \notin A_i$ so that $x \notin \bigcup_{i \in \mathbb{N}} A_i$. As U_0 was chosen arbitrary we have proved that $(\bigcup_{i \in \mathbb{N}} A_i)^\circ = \emptyset$ for if $y \in (\bigcup_{i \in \mathbb{N}} A_i)^\circ$ then $(\bigcup_{i \in \mathbb{N}} A_i)^\circ \in \mathcal{A}$ and we have then that $y \notin \bigcup_{i \in \mathbb{N}} A_i$ contradicting the fact that $y \in (\bigcup_{i \in \mathbb{N}} A_i)^\circ \subseteq \bigcup_{i \in \mathbb{N}} A_i \Rightarrow y \in \bigcup_{i \in \mathbb{N}} A_i$. □

We use the Baire category theorem to prove the open mapping theorem.

Theorem 12.408. (Open Mapping Theorem) *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be Banach spaces over $\mathbb{K}$ and $L: X \rightarrow Y$ a surjective continuous linear function then L is also a open function.*

Proof. First if $\|L\| = 0$ then we have $\forall x \in X$ that $\|L(x)\|_Y \leqslant \|L\| \cdot \|x\| = 0 \Rightarrow L(x) = 0$ so as L is a surjection we have that $Y = \{0\}$ and thus also $X = \{0\}$ so that the only open set in $X$is $\{0\}$ and $L(\{0\}) = \{0\} = Y$ is open proving that L is a open function. Let's now take the case that $\|L\| > 0$. We split the proof in little sub-proofs.

1. First we show that given $\delta > 0$ and $B_{\|\|_X}(0, \delta)$ a open ball around 0 then there exists $\lambda > 0$ with $B_{\|\|_Y}(0, \lambda) \subseteq \overline{L(B_{\|\|_X}(0, \delta))}$

   Proof. Take $\delta > 0$ and $x \in X$ then given the Archimedean order property of $\mathbb{R}$ (see 9.54) there exists a $n \in \mathbb{N}$ such that $\frac{2}{\delta} \cdot \|x\| < n \cdot 1 = n$ so that $\|x\| < \frac{n \cdot \delta}{2}$ so if $y = \frac{x}{n} \Rightarrow x = y \cdot n \Rightarrow \|n \cdot y\| < \frac{n \cdot \delta}{2} \Rightarrow \|y\| < \frac{\delta}{2} \Rightarrow x \in n \cdot B_{\|\|_X}\left(0, \frac{\delta}{2}\right) \Rightarrow x \in \bigcup_{n \in \mathbb{N}} \left(n \cdot B_{\|\|_X}\left(0, \frac{\delta}{2}\right)\right)$ proving that

   $$X = \bigcup_{n \in \mathbb{N}} n \cdot B_{\|\|_X}\left(0, \frac{\delta}{2}\right) \tag{12.82}$$

As L is surjective we have that $Y = L(X) = \bigcup_{n\in\mathbb{N}} L\left(n\cdot B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right) = \bigcup_{n\in\mathbb{N}} n\cdot L\left(B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right) \subseteq \bigcup_{n\in\mathbb{N}} \overline{n\cdot L\left(B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right)} \subseteq Y$ giving

$$Y = \bigcup_{n\in\mathbb{N}} \overline{n\cdot L\left(B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right)} \tag{12.83}$$

Note that that $0\in Y = Y^\circ \underset{12.83}{=} \left(\bigcup_{n\in\mathbb{N}} \overline{n\cdot L\left(B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right)}\right) \Rightarrow \bigcup_{n\in\mathbb{N}} \overline{n\cdot L\left(B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right)} \neq \emptyset$ we must as Y is a Banach space (thus a complete metric space) using the Baire Category Theorem (12.407) there must exists a $n\in\mathbb{N}$ such that $\left(\overline{n\cdot L\left(B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right)}\right)^\circ \neq \emptyset$. So $\exists z\in\left(\overline{n\cdot L\left(B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right)}\right)^\circ$ which is open so there exists a $\beta>0$ such that $z\in B_{\|\|_Y}(z,\beta)\subseteq\left(\overline{n\cdot L\left(B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right)}\right)^\circ \subseteq \overline{n\cdot L\left(B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right)} \underset{12.76}{=} n\cdot\overline{L\left(B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right)}$ so that

$$\frac{1}{n}\cdot B_{\|\|_Y}(z,\beta)\subseteq \overline{L\left(B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right)} \tag{12.84}$$

As $B_{\|\|_X}\left(0,\frac{\delta}{2}\right) - B_{\|\|_X}\left(0,\frac{\delta}{2}\right) \subseteq B_{\|\|_Y}(0,\delta)$ [If $x\in B_{\|\|_X}\left(0,\frac{\delta}{2}\right) - B_{\|\|_X}\left(0,\frac{\delta}{2}\right) \Rightarrow x = x_! = x_2$ with $\|x_1\|_X < \frac{\delta}{2}, \|x_2\|_X < \frac{\delta}{2} \Rightarrow \|x\|_X \leqslant \|x_1\|_X + \|x_2\|_X < \frac{\delta}{2}+\frac{\delta}{2} = \delta$] we have $L\left(B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right) - L\left(B_{\|_X\|}\left(0,\frac{\delta}{2}\right)\right) = L\left(B_{\|\|_X}\left(0,\frac{\delta}{2}\right) - B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right) = L\left(B_{\|\|_X}\left(0,\frac{\delta}{2}\right) - B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right) \subseteq L(B_{\|\|_X}(0,\delta))$ giving

$$L\left(B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right) - L\left(B_{\|_X\|}\left(0,\frac{\delta}{2}\right)\right) \subseteq L(B_{\|\|_X}(0,\delta)) \tag{12.85}$$

Further $\bigcup_{x\in\frac{1}{n}B_{\|\|_Y}(z,\beta)}\left(x - \frac{1}{n}B_{\|\|_Y}(z,\beta)\right) \subseteq \frac{1}{n}B_{\|\|_Y}(z,\beta) - \frac{1}{n}B_{\|\|_Y}(z,\beta) \subseteq_{12.84} \overline{L\left(B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right)} - \overline{L\left(B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right)} \subseteq_{12.76} \overline{L\left(B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right) - L\left(B_{\|\|_X}\left(0,\frac{\delta}{2}\right)\right)} \subseteq_{12.85} \overline{L(B_{\|\|_X}(0,\delta))}$ so that we have for $U = \bigcup_{x\in\frac{1}{n}B_{\|\|_Y}(z,\beta)}\left(x-\frac{1}{n}B_{\|\|_Y}(z,\beta)\right)$

$$U = \bigcup_{x\in\frac{1}{n}B_{\|\|_Y}(z,\beta)}\left(x-\frac{1}{n}B_{\|\|_Y}(z,\beta)\right)\subseteq\overline{L(B_{\|\|_X}(0,\delta))} \tag{12.86}$$

Now using 12.76 again we have that $x-\frac{1}{n}\cdot B_{\|\|_Y}(z,\beta)$ is open so that U being the union of open sets must be open, as $0 = z - z\in U$ there exists a $\lambda>0$ such that $0\in B_{\|\|_Y}(0,\lambda)\subseteq U$ or using 12.86

$$B_{\|\|_Y}(0,\lambda)\subseteq\overline{L(B_{\|\|_X}(0,\delta))} \tag{12.87}$$

what we set out to prove. □

2. Next we refine (1) to show that

$$\text{If } \sigma_0 > 0 \text{ then } \exists \lambda > 0 \text{ such that } B_{\|\|_Y}(0, \lambda) \subseteq L(B_{\|\|_X}(0, 2 \cdot \sigma_0)) \tag{12.88}$$

Proof. Using the Archimedean property of the reals (see 9.55) we find a $N \in \mathbb{N}$ such that $\frac{1}{N} < \sigma_0$. Define then $\forall n \in \mathbb{N}$ $\sigma_n = \left(\frac{1}{N+1}\right)^n$, we have then that $\sum_{i=1}^{\infty} \sigma^i = \sum_{i=1}^{\infty} \left(\frac{1}{N+!}\right)^i \underset{\frac{1}{N+1} < 1 \text{ and } 12.403}{=} \frac{\frac{1}{N+1}}{1 - \frac{1}{N+!}} = \frac{\frac{1}{N+1}}{\frac{N+1-1}{N+1}} = \frac{\frac{1}{N+1}}{\frac{N}{N+1}} = \frac{1}{N} < \sigma_0$ or

$$\sum_{i=1}^{\infty} \sigma_i = \frac{1}{N} < \sigma_0 \tag{12.89}$$

Now $\forall n \in \mathbb{N}_0 = \mathbb{N} \bigcup \{0\}$ we have using (1) the existence of a $\lambda'_n > 0$ such that $B_{\|\|_Y}(0, \lambda'_n) \subseteq \overline{L(B_{\|\|_X}(0, \sigma_n))}$. Define $0 < \lambda_n = \min\left(\lambda'_n, \frac{1}{n+1}\right) < \frac{1}{n+1}$ then trivially we have $B_{\|\|_Y}(0, \lambda_n) \subseteq \overline{L(B_{\|\|_X}(0, \sigma_n))}$ and using 12.319 and 12.338 we have that $\lim_{n \to \infty} \lambda_n = 0$. So we have found a sequence $\{\lambda_i\}_{i \in \mathbb{N}_0}$ such that

$$\lim_{i \to \infty} \lambda_i \wedge \forall n \in \mathbb{N}_0 \text{ we have } B_{\|\|_Y}(0, \lambda_n) \subseteq \overline{L(B_{\|\|_X}(0, \sigma_n))} \tag{12.90}$$

Taken now $y \in B_{\|\|_Y}(0, \sigma_0)$. Then $y \in B_{\|\|_Y}(0, \lambda_0) \Rightarrow y \in \overline{L(B_{\|\|_X}(0, \sigma_0))}$ so there exists a $y_0 \in B_{\|\|_Y}(0, \lambda_1) \bigcap L(B_{\|\|_X}(0, \sigma_0))$ [as $y \in B_{\|\|_Y}(0, \lambda_1)$ a open set] thus $\exists x_0 \in B_{\|\|_X}(0, \sigma_0)$ such that $y_0 = L(x_0) \in B_{\|\|_Y}(0, \lambda_1)$. This gives

$$\exists x_0 \in B_{\|\|_X}(0, \sigma_0) \text{ with } \|y - L(x_0)\|_Y < \lambda_1 \tag{12.91}$$

Take now $\mathcal{N}_0 = \{\{x_i\}_{i \in \{0, \ldots, n\}} | n \in \mathbb{N}_0 \wedge \forall i \in \{0, \ldots, n\}$ we have $x_i \in B_{\|\|_X}(0, \sigma_i) \wedge \|y - L(\sum_{j=0}^{i} x_j)\|_Y < \lambda_{i+1}\}$ then we have

$$\{x_0\}_{i \in \{0, \ldots, n\}} \in \mathcal{N}_0 \tag{12.92}$$

If $\{x_i\}_{i \in \{1, \ldots, n\}} \in \mathcal{N}_0$ then as $\|y - L(\sum_{j=0}^{n} x_i)\|_Y < \lambda_{n+1}$ we have that $y - L(\sum_{j=0}^{n} x_j) \in B_{\|\|_Y}(0, \lambda_{n+1}) \subseteq \overline{L(B_{\|\|_X}(0, \sigma_{n+1}))}$ so that there exists a $y_{n+1} \in B_{\|\|_Y}(y - L(\sum_{j=0}^{n} x_j), \lambda_{n+2}) \bigcap L(B_{\|\|_X}(0, \sigma_{n+1})) \Rightarrow \exists x_{n+1} \in B_{\|\|_X}(0, \sigma_{n+1})$ such that $y_{n+1} = L(x_{n+1}) \in B_{\|\|_Y}(y - L(\sum_{j=0}^{n} x_j), \lambda_{n+2})$. Then $\|y - L(\sum_{j=0}^{n+1} x_j)\|_Y = \|(y - L(\sum_{j=0}^{n})) - L(x_{n+1})\|_Y < \lambda_{n+1} = \lambda_{(n+1)+1}$. This defines a function $\rho: \mathcal{N}_0 \to X$ such that $x_{n+1} = \rho(\{x_i\}_{i \in \{0, \ldots, n\}})$ and $\{x_i\}_{i \in \{0, \ldots, n+1\}} \in \mathcal{N}_0$. Using recursion (see 5.26 this defines a sequence (or function from $\mathbb{N}_0 \to X$) $\{x_i\}_{i \in \mathbb{N}_0}$ such that

$$\forall i \in \mathbb{N}_0 \text{ we have } x_i \in B_{\|\|_X}(0, \sigma_i) \wedge \left\| y - L\left(\sum_{j=0}^{i} x_i\right) \right\|_Y < \lambda_{i+1} \tag{12.93}$$

As $x_i \in B_{\|\|_X}(0, \sigma_i)$ we have $\|x_i\|_X < \sigma_i$ so that using 12.381 we have that $\sum_{i=0}^{\infty} x_i$ converges and $\|\sum_{i=0}^{\infty} x_i\|_X \leqslant \sum_{i=0}^{\infty} \sigma_i \underset{12.367}{=} \sigma_0 + \sum_{i=1}^{\infty} \sigma_i <_{12.89} \sigma + \sigma = 2 \cdot \sigma$ so if $x = \sum_{i=0}^{\infty} x_i$ then $\|x\|_X < 2 \cdot \sigma$ and thus

$$x = \sum_{j=0}^{\infty} x_i \in B_{\|\|_X}(0, 2 \cdot \sigma_0) \tag{12.94}$$

Assume now that $\|y - L(x)\|_Y > 0$ then we can take $\varepsilon = \|y - L(x)\|_Y > 0$ so that as $\lim_{i \rightarrow \infty} \lambda_i = 0$ and $0 < \lambda_i$ that there exists a $N_1 \in \mathbb{N}_0$ such that if $n \geqslant N_1$ then $\lambda_n = |\lambda_n - 0| < \frac{\varepsilon}{2}$. As $x = \sum_{i=0}^{\infty} x_i$ there exists a $N_2 \in \mathbb{N}_0$ such that if $n \geqslant N_2$ then $\|x - \sum_{i=0}^{n} x_i\|_X < \frac{\varepsilon}{2 \cdot \|L\|}$. If $n \geqslant \max(N_1, N_2)$ then $\varepsilon = \|y - L(x)\|_Y = \|y - L(\sum_{i=0}^{n} x_i) + L(\sum_{i=0}^{n} x_i) - L(x)\|_Y \leqslant \|y - L(\sum_{i=0}^{n} x_i)\|_Y + \|L(\sum_{i=0}^{n} x_i) - L(x)\|_Y = \|y - L(\sum_{i=0}^{n} x_i)\|_Y + \|L(\sum_{i=0}^{n} x_i - x)\|_Y < \lambda_{n+1} + \|L\| \cdot \|\sum_{i=0}^{n} x_i - x\| < \frac{\varepsilon}{2} + \|L\| \cdot \frac{\varepsilon}{2 \cdot \|L\|} = \varepsilon$ giving the contradiction $\varepsilon < \varepsilon$. So we must have that $\|y - L(x)\| = 0 \Rightarrow y = L(x) \underset{12.94}{\Rightarrow} y \in L(B_{\|\|_X}(0, 2 \cdot \sigma_0))$. As $y \in B_{\|\|_Y}(0, \sigma_0)$ was chosen arbitrary we have that $B_{\|\|_Y}(0, \sigma_0) \subseteq L(B_{\|\|_X}(0, 2 \cdot \sigma_0))$, which taking $\lambda = \sigma_0$ proves our assertion.

3. Now take $B_{\|\|_X}(0, 1) = B_{\|\|_X}\left(0, 2 \cdot \frac{1}{2}\right)$ we find using (2) the existence of a $\lambda > 0$ such that $B_{\|\|_Y}(0, \lambda) \subseteq L(B_{\|\|_X}(0, 1))$ which using 12.187 means that L is open.

□ □

Corollary 12.409. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be Banach spaces over $\mathbb{K}$ and $L: X \rightarrow Y$ a continuous linear isomorphism then L^{-1} is continuous.*

Proof. This is trivial using the open function theorem (12.408) and 12.141. □

12.12 Integration in Banach space

Note 12.410. In the following we assume that with $[a, b]$ we mean a closed and bounded interval in $\mathbb{R}$ that is not empty (or $a < b$)

Definition 12.411. $\mathbb{N}_{0\,1} = \{1, \ldots\} = \mathbb{N} \setminus \{1\}$

Definition 12.412. *A partition $\mathcal{P}$ of $[a, b]$ is a family $\{t_i\}_{i \in \{1, \ldots, n\}}$ $n \in \mathbb{N}_{0\,1}$ such that $t_1 = a, t_n = b$ and $\forall i \in \{1, \ldots, n-1\}$ we have $t_i < t_{i+1}$.*

We prove now some trivial facts for a partition of a interval

Theorem 12.413. *Let $[a, b] \subseteq \mathbb{R}$ and $\mathcal{P} = \{t_i\}_{i \in \{1, \ldots, n\}}$ a partition of $[a, b]$ then $[a, b] = \bigcup_{i \in \{1, \ldots, n-1\}} [t_i, t_{i+1}]$*

Proof. As $\forall i \in \{1, \ldots, n\}$ we have $t_i \in [a, b] \Rightarrow a \leqslant t_i \leqslant b$ so that $\forall i \in \{1, \ldots, n-1\}$ we have $a \leqslant t_i < t_{i+1} \leqslant b \Rightarrow [t_i, t_{i+1}] \subseteq [a, b] \Rightarrow \left(\bigcup_{i \in \{1, \ldots, n-1\}} [t_i, t_{i+1}]\right) \subseteq [a, b]$. Take now $x \in [a, b]$ and take $M_x = \{i \in \{1, \ldots, n-1\} | t_i \leqslant x\}$ which is finite and as $t_1 = a \leqslant x \Rightarrow 1 \in M_x$ is not empty, so $m = \max(M_x)$ exist. As $m \in M_x$ we have $t_m \leqslant x$, if $t_{m+1} < x \Rightarrow t_{m+1} \leqslant x \Rightarrow m+1 \in M_x \Rightarrow m+1 \leqslant m$ a contradiction so $x \leqslant t_{m+1}$. So $x \in [t_m, t_{m+1}] \subseteq \bigcup_{i \in \{1, \ldots, n-1\}} [t_i, t_{i+1}] \Rightarrow [a, b] \subseteq \bigcup_{i \in \{1, \ldots, n-1\}} [t_i, t_{i+1}]$ □

Lemma 12.414. *Let $\{t_i\}_{i \in \{1, \ldots, n\}}$, $n \in$ be such that $\forall i \in \{1, \ldots, n-1\}$ we have $t_i < t_{i+1}$ then*

1. *$\forall i, j \in \{1, \ldots, n\}$ with $i < j$ we have $t_i < t_j$*
2. *If $t_i = t_j$ then $i = j$*
3. *If $i \in \{1, \ldots, n-1\}$, $k \in \{1, \ldots, n\}$ and $t_k \in [t_i, t_{i+1}]$ then $k = i$ or $k = i+1$*

Proof.

1. We prove this by induction, so given $i \in \{1, \ldots, n\}$ let $\mathcal{S}_i = \{k \in \mathbb{N} | \text{if } i + k \leqslant n \text{ then } t_i < t_{i+k}\}$ then:
   a. If $k = 1$ then if $i + k = i + 1 \leqslant n$ we have $i \in \{1, \ldots, n-1\}$ and by the hypothesis we have $t_i < t_{i+1} = t_{i+k}$ so $1 \in \mathcal{S}_i$
   b. If $k \in \mathcal{S}_i$ then for $k+1$ we have if $i + (k+1) \leqslant n \rightarrow (i+k) + 1 \leqslant n$ that $t_{i+k} < t_{(i+k)+1}$ (hypothesis) and $t_i < t_{i+k}$ ($k \in S_i$) so that $t_i < t_{(i+k)+1} = t_{i+(k+1)}$ proving that $k+1 \in \mathcal{S}_i$

   Using induction we have that $\mathcal{S}_i = \mathbb{N}$ so if $i, j \in \{1, \ldots, n\}$ with $i < j$ then $k = j - i \in \mathbb{N} \in \mathcal{S}_i$ and $i + k = j \leqslant n$ so that $t_i < t_{i+k} = t_j$.
2. If $t_i = t_k$ then if $i < k$ we have by (1) $t_i < t_k$ a contradiction, if $k < i$ we have by (1) $t_k < t_i$ a contradiction. So the only possibility left is $i = k$
3. If $t_k \in [t_i, t_{i+1}]$ then $t_i \leqslant t_k \leqslant t_{i+1}$. If now $k < i$ then using (1) we have $t_k < t_i$ contradicting $t_i \leqslant t_k$, if $i + 1 < k$ we have using (1) $t_{i+1} < t_k$ contradicting $t_k \leqslant t_{i+1}$. So we must have $i \leqslant k \leqslant i+1$ which means $i = k$ or $k = i+1$ □

Theorem 12.415. *Let $\mathcal{P}_1 = \{t_i^1\}_{i \in \{1, \ldots, n_1\}}$, $\mathcal{P}_2 = \{t_i^2\}_{i \in \{1, \ldots, n_2\}}$ be two partitions of $[a, b]$ then there exists a unique partition $\mathcal{P} = \{t_i\}_{i \in \{1, \ldots, n\}}$ of $[a, b]$ such that $\{t_i | i \in \{1, \ldots, n\} |\} = \{t_i^1 | i \in \{1, \ldots, n_1\}\} \bigcup \{t_i^2 | i \in \{1, \ldots, n_2\}\}$. This partition of $[a, b]$ is noted by $\mathcal{P}_1 \boxplus \mathcal{P}_2$.*

Proof. Define $P = \{t_i^1 | i \in \{1, \ldots, n_1\}\} \bigcup \{t_i^2 | i \in \{1, \ldots, n_2\}\} \subseteq [a, b]$ which is a finite set. Using 5.53 there exists a unique bijection $t: \{1, \ldots, \#(P)\} \rightarrow P$ such that $\forall i \in \{1, \ldots, \#(P) - 1\}$ we have $t(i) < t(i+1)$. As $a, b \in P$ there exists a $i, j \in \{1, \ldots, \#(P)\}$ such that $t(i) = a$ and $t(j) = b$. If $i > 1$ then $a = t(i) > t(1) \geqslant a$ a contradiction so we must have $i = 1$. If $j < \#(P)$ then $b = t(j) < t(\#(P)) \leqslant b$ a contradiction so we must have $j = \#(P)$. If we then take $n = \#(P)$ and $\mathcal{P} = \{t(i)\}_{i \in \{1, \ldots, n\}}$ then we have that $t(1) = a, t(n) = b$ and $\forall i \in \{1, \ldots, n-1\}$ that $t(i) = t(i+1)$ proving our theorem. □

Lemma 12.416. *Let $\mathcal{P}_1 = \{t_i^1\}_{i \in \{1, \ldots, n_1\}}$, $\mathcal{P}_2 = \{t_i^2\}_{i \in \{1, \ldots, n_2\}}$ be two partitions of $[a, b]$ with $\mathcal{P} = \mathcal{P}_1 \boxplus \mathcal{P}_2 = \{t_i\}_{i \in \{1, \ldots, n\}}$ then for every $i \in \{1, \ldots, n-1\}$ there exists unique $i_1 \in \{1, \ldots, n_1 - 1\}, i_2 \in \{1, \ldots, n_2 - 1\}$ such that $[t_i, t_{i+1}] \subseteq [t_{i_1}^1, t_{i_1+1}^1]$ and $[t_i, t_{i+1}] \subseteq [t_{i_2}^2, t_{i_2+1}^2]$.*

Proof. We prove this for i_1 the proof for i_2 is similar (by switching $1 \leftrightarrow 2$). First as $i \in \{1, \ldots, n-1\}$ we have $t_i \neq b = t_{n_1} = t_{n_2}$, then as $t_i \in \{t_j | j \in \{1, \ldots, n\}|\} = \{t_j^1 | j \in \{1, \ldots, n_1\}\} \bigcup \{t_j^2 | j \in \{1, \ldots, n_2\}\}$ we have either

1. $(\boldsymbol{t_1 \in \{t_j^1 | j \in \{1, \ldots, n_1 - 1\}\}})$ then there exists a $k \in \{1, \ldots, n_1 - 1\}$ such that $t_i = t_k^1$, by the definition of $\mathcal{P} = \mathcal{P}_1 \boxplus \mathcal{P}_2$ there exists a $l \in \{1, \ldots, n\}$ such that $t_l = t_{k+1}^1$. Suppose now that $t_{k+1}^1 < t_{i+1} \Rightarrow t_l < t_{i+1}$ then $t_i = t_k^1 < t_{k+1}^1 = t_l < t_{i+1} \Rightarrow t_l \in [t_i, t_{i+1}] \underset{12.414}{\Rightarrow} l = i$ or $l = i+1$, as $t_i < t_l$ we must have $i \neq l$ so that $l = i+1$ but this contract $t_l < t_{i+1}$. The conclusion is thus that $t_{i+1} \leqslant t_{k+1}^1 \Rightarrow t_k^1 = t_i < t_{i+1} \leqslant t_{k+1}^1 \Rightarrow [t_i, t_{i+1}] \subseteq [t_k^1, t_{k+1}^1]$ so we take $i_1 = k$ to get $[t_i, t_{i+1}] \subseteq [t_{i_1}^1, t_{i_1+1}^1]$.

2. $(\boldsymbol{t_1 \in \{t_j^2 | j \in \{1, \ldots, n_2 - 1\}\}})$ then there exists a $k \in \{1, \ldots, n_2 - 1\}$ such that $t_k^2 = t_i$. Define then $B = \{j \in \{1, \ldots, n_1\} | t_j^1 \leqslant t_i\}$ then as $t_1^1 = a \leqslant a = t_1$ we have $1 \in B \Rightarrow B \neq 0$ and as B is finite (subset of a finite set) there exist a $m = \max(B) \in \{1, \ldots, n_1\}$. If $m = n_1$ then $b = t_m^1 = t_{n_1}^1 = b \leqslant t_i < b$ a contradiction so $m \in \{1, \ldots, n_1 - 1\}$. Then we have by the definition of a maximum and B that $t_m^1 \leqslant t_i < t_{m+1}^1$. Assume now that $t_{m+1}^1 < t_{i+1}$ then as by the definition of $\mathcal{P}$ there exists a $l \in \{1, \ldots, n\}$ such that $t_l = t_{m+1}^1$ we have from $t_i < t_{m+1}^1 = t_l < t_{i+1} \Rightarrow t_l \in [t_i, t_{i+1}] \underset{12.414}{\Rightarrow} l = i$ or $l = i+1$ which is impossible as $t_i < t_l < t_{i+1}$, we are thus forced to conclude that $t_{i+1} \leqslant t_{m+1}^1$. This gives $t_m^1 \leqslant t_i < t_{i+1} \leqslant t_{m+1}^1 \Rightarrow [t_i, t_{i+1}] \subseteq [t_m^1, t_{m+1}^1]$ giving if we take $i_1 = m$ that $[t_i, t_{i+1}] \subseteq [t_m^1, t_{m+1}^1]$

Next we prove uniqueness so assume that there exists a i_1, i_1' such that $[t_i, t_{i+1}] \subseteq [t_{i_1}^1, t_{i_1+1}^1], [t_{i_1'}^1, t_{i_1'+1}^1]$ and assume that $i_1 \neq i_1'$ we have then two cases

1. $(\boldsymbol{i_1 < i_1'})$ then $t_{i_1}^1 < t_{i_1'}^1$ and so $t_{i_1}^1 < t_{i_1'}^1 \leqslant t_i < t_{i+1} \leqslant t_{i_1+1}^1 \Rightarrow t_{i_1'}^1 \in [t_{i_1}^1, t_{i_1+1}^1] \underset{12.414}{\Rightarrow} i_1' = i_1$ or $i_1' = i_1 + 1$ which as $t_{i_1}^1 < t_{i_1'}^1$ means that $i_1' = i_1 + 1$ but then $t_{i_1+1}^1 = t_{i_1'}^1 \leqslant t_i < t_{i_1+1}^1$ a contradiction.

2. $(\boldsymbol{i_1' < i_1})$ then $t_{i_1'}^1 < t_{i_1}^1$ and so $t_{i_1'}^1 < t_{i_1}^1 \leqslant t_i < t_{i+1} \leqslant t_{i_1'+1}^1 \Rightarrow t_{i_1}^1 \in [t_{i_1'}^1, t_{i_1'+1}^1] \underset{12.414}{\Rightarrow} i_1 = i_1'$ or $i_1 = i_1' + 1$ which as $t_{i_1'}^1 < t_{i_1}^1$ means that $i_1 = i_1' + 1$ but then $t_{i_1'+1}^1 = t_{i_1}^1 \leqslant t_i < t_{i_1'+1}^1$ a contradiction.

so we must conclude that $i_1 = i_1'$ □

Lemma 12.417. *Let $\mathcal{P}_1 = \{t_i^1\}_{i \in \{1, \ldots, n_1\}}$, $\mathcal{P}_2 = \{t_i^2\}_{i \in \{1, \ldots, n_2\}}$ be partitions of $[a, b]$ and $\mathcal{P} = \mathcal{P}_1 \boxplus \mathcal{P}_2 = \{t_i\}_{i, \{1, \ldots, n\}}$ then using the previous lemma for every $i \in \{0, \ldots, n-1\}$ there exists unique $i_1 \in \{1, \ldots, n_1 - 1\}, i_2 \in \{1, \ldots, n_2 - 1\}$ so that $[t_i, t_{i+1}] \subseteq [t_{i_1}^1, t_{i_1-1}^1], [t_{i_2}^2, t_{i_2+1}^2]$. This defines functions $i_1: \{1, ., n-1\} \rightarrow \{1, \ldots, n_1 - 1\}$, $i_2: \{1, \ldots, n-1\} \rightarrow \{1, \ldots, n_2 - 1\}$ such that $[t_i, t_{i+1}] \subseteq [t_{i_1(i)}^1, t_{i_2(i)+1}^1], [t_{i_2(i)}^2, t_{i_2(i)+1}^2]$. We have then additional that*

1. *i_1, i_2 are surjective*
2. *$\forall i \in \{1, \ldots, n_1 - 1\}$ we have that*
   a. *$i_1^{-1}(\{i\}) = \{m_i^1, M_i^1\}$ where $t_{m_i^1} = t_i^1$ and $t_{M_i^1+1} = t_{i+1}^1$*

b. $i_2^{-1}(\{i\}) = \{m_i^2, M_i^2\}$ *where* $t_{m_i^2} = t_i^2$ *and* $t_{M_i^2+1} = t_{i+1}^2$

3. *If* $i \neq j$ *then* $i_1^{-1}(\{i\}) \bigcap i_1^{-1}(\{j\}) = \emptyset$ *and* $i_2^{-1}(\{i\}) \bigcap i_2^{-1}(\{j\}) = \emptyset$

Proof. We prove this for i_1 (the proof for i_2 is similar by interchanging $1 \leftrightarrow 2$)

1. Let $k \in \{1, ..., n_1 - 1\}$ then for t_k^1 there exists a $i \in \{1, ..., n\}$ such that $t_k^1 = t_i$ and as $k < n_1 \Rightarrow t_i = t_k^1 < t_{n_1}^1 = b = t_n$ we must have $i \in \{1, ..., n-1\}$. Now assume that $t_{k+1}^1 < t_{i+1}$ then as there exists a $j \in \{1, ..., n\}$ such that $t_j = t_{k+1}^1$ we have then $t_i = t_k^1 < t_{k+1}^1 = t_j < t_{i+1} \Rightarrow t_j \in [t_i, t_{i+1}] \underset{12.414 \text{ and } t_i < t_j}{\Rightarrow} j = i+1$ but then $t_{i+1} = t_j = t_{k+1}^1 < t_{i+1}$ a contradiction. So we must have $t_{i+1} \leqslant t_{k+1}^1$ or $t_k^1 = t_i < t_{i+1} \leqslant t_{k+1}^1 \Rightarrow [t_i, t_{i+1}] \subseteq [t_k^1, t_{k+1}^1]$ our $i_1(i) = k$ proving surjectivity.

2. Let $i \in \{1, ..., n_1 - 1\}$, as i_1 is surjective we have that $\emptyset \neq i_1^{-1}(\{i\}) \subseteq \{1, ..., n-1\}$ so there exists $m_i^1 = \min(i_1^{-1}(\{i\}))$, $M_i^1 = \max(i_1^{-1}(\{i\}))$. We prove now that $i_1^{-1}(\{i\}) = \{m_i^1, M_i^1\}$

   a. $i_1^{-1}(\{i\}) \subseteq \{m_i^1, M_i^1\}$ (this is trivial by the definition of minimum and maximum)

   b. If $k \in \{m_i^1, M_i^1\}$ then $m_i^1 \leqslant k \leqslant M_i^1 \Rightarrow t_{m_i} \leqslant t_k$ and $k+1 \leqslant M_i^1 + 1 \Rightarrow t_{m_i^1+1} \leqslant t_{M_i^1+1}$. As we have $[t_{m_i^1}, t_{m_i^1+1}], [t_{M_i^1}, t_{M_i^1+1}] \subseteq [t_i^1, t_{i+1}^1]$ we have $t_i^1 \leqslant t_{m_i^1} \leqslant t_k < t_{k+1} \leqslant t_{M_i^1+1}^1 \leqslant t_{i+1}^1 \Rightarrow [t_k, t_{k+1}] \subseteq [t_i^1, t_{i+1}^1] \Rightarrow i_1(k) = i \Rightarrow k \in i_1^{-1}(\{i\})$. Proving $\{m_i^1, M_i^1\} \subseteq i_1^{-1}(\{i\})$

   As $i_1^{-1}(\{i\}) = \{m_i^1, ..., M_i^1\}$ we have $m_i^1, M_i^1 \in i_1^{-1}(\{i\}) \Rightarrow i_1(m_i^1) = i, i_1(M_i^1) = i$ or

   $$[t_{m_i^1}, t_{m_i^1+1}], [t_{M_i^1}, t_{M_i^1+1}] \subseteq [t_i^1, t_{i+1}^1] \tag{12.95}$$

   From 12.95 it follows that $t_i^1 \leqslant t_{m_i^1}$, assume now that $t_i^1 < t_{m_i^1}$ then by the definition of $\mathcal{P}_1 \boxplus \mathcal{P}_2$ there exists a $l \in \{1, ..., n\}$ such that $t_l = t_i^1$ then $t_l = t_1^1 < t_{m_i^1}$ so that $l < m_i^1$ or $l + 1 < m_i^1 + 1 \Rightarrow t_1^1 = t_l < t_{l+1} < t_{m_i^1+1} \leqslant t_{i+1}^1 \Rightarrow [t_l, t_{l+1}] \subseteq [t_i^1, t_{i+1}^1]$ proving that $i_1(l) = i \Rightarrow l \in i_1^{-1}(\{i\}) = \{m_i^1, ..., M_i^1\}$ proving that $m_i^1 \leqslant l < m_i^1$ a contradiction. So we must have $t_i^1 = t_{m_i^1.}$ From 12.95 it follows that $t_{M_i^1+1} \leqslant t_{i+1}^1$, assume now that $t_{M_i^1+1} < t_{i+1}^1$ then by the definition of $\mathcal{P}_1 \boxplus \mathcal{P}_2$ there exists a $l \in \{1, ..., n\}$ so that $t_{i+1}^1 = t_l$ then $t_{M_i^1+1} < t_{i+1}^1 = t_l$ so that $M_i^1 + 1 < l \Rightarrow M_i^1 < l - 1$ thus $t_i^1 \leqslant t_{M_i^1} < t_{l-1} < t_l = t_{i+1}^1 \Rightarrow [t_{l-1}, t_l] \subseteq [t_i^1, t_{i+1}^1]$ proving that $i_1(l-1) = i \Rightarrow l - 1 \in i_1^{-1}(\{i\}) = \{m_i^1, ..., M_i^1\} \Rightarrow l - 1 \leqslant M_i^1 < l - 1$ a contradiction. So we must have that $t_{M_i^1+1} = t_{i+1}^1$

3. Assume that $k \in i_1^{-1}(i) \bigcap i_1^{-1}(j), i \neq j$ then $i_1(k) = i$ and $i_1(k) = j$ so that by definition of a function we must have $i = j$. $\square$

Definition 12.418. *Let* $[a, b] \subseteq \mathbb{R}$ *and* $\mathcal{P} = \{t_i\}_{i \in \{1, ..., n\}}$ *be a partition of* $[a, b]$ *then the norm* $\mu(\mathcal{P})$ *of the partition is defined by* $\mu(\mathcal{P}) = \max(\{|t_{i+1} - t_i| | i \in \{1, ..., n-1\}\}) \underset{t_i < t_{i+1}}{=} \max(\{t_{i+1} - t_i | i \in \{1, ..., n-1\}\})$

Definition 12.419. *Let $[a,b] \subseteq \mathbb{R}$ and $\mathcal{P} = \{t_i\}_{i \in \{1,\ldots,n\}}$ be a partition of $[a,b]$ then a **tag** on a partition is a family $\{s_i\}_{i \in \{1,\ldots,n-1\}}$ such that $\forall i \in \{1,\ldots,n-1\}$ we have $s_i \in [t_i, t_{i+1}]$. A tagged partition $\mathbb{P}$ is a pair of a partition and a tag on this partition, so $\mathbb{P} = (\mathcal{P}, s) = (\{t_i\}_{i \in \{1,\ldots,n\}}, \{s_i\}_{i \in \{1,\ldots,n-1\}})$ such that $\forall i \in \{1,\ldots,n\}$ we have $t_i < t_{i+1} \wedge s_i \in [t_i, t_{i+1}]$ and $t_i = a, t_n = b$. The norm of a tagged partition is the norm of its partition so $\mu(\mathbb{P}) = \mu((\mathcal{P}, \{s_i\}_{i \in \{1,\ldots,n-1\}})) = \mu(\mathcal{P})$*

Definition 12.420. *Let $[a,b] \subseteq \mathbb{R}$, $\mathbb{P} = \{\{t_i\}_{i \in \{1,\ldots,n\}}, \{s_i\}_{i \in \{1,\ldots,n-1\}}\}$ is a tagged partition on $[a,b]$, $\langle X, \|\|\rangle$ a normed real vector space and $f: [a,b] \rightarrow X$ a function then a Riemann sum of f using the tagged partition, noted by $\mathcal{S}(f, \mathbb{P})$ is defined by $\mathcal{S}(f, \mathbb{P}) = \sum_{i=1}^{n-1} f(s_i) \cdot (t_{i+1} - t_i)$.*

Theorem 12.421. *Let $[a,b] \subseteq \mathbb{R}$, $\mathbb{P} = \{\{t_i\}_{i \in \{1,\ldots,n\}}, \{s_i\}_{i \in \{1,\ldots,n-1\}}\}$ a tagged partition on $[a,b]$, $f, g: [a,b] \rightarrow \mathbb{R}$ a function such that $\forall x \in [a,b]$ we have $f(x) \leqslant g(x)$ then $\mathcal{S}(f, \mathbb{P}) \leqslant \mathcal{S}(g, \mathbb{P})$*

Proof. As $\forall i \in \{1,\ldots,n\}$ we have $0 \leqslant t_{i+1} - t_i$, $f(s_i) \leqslant g(s_i)$ it follows that $f(s_i) \cdot (t_{i+1} - t_i) \leqslant g(s_i) \cdot (t_{i+1} - t_i)$ so that $\mathcal{S}(f, \mathbb{P}) = \sum_{i=1}^{n-1} f(s_i) \cdot (t_{i+1} - t_i) = \sum_{i=1}^{n-1} g(s_i) \cdot (t_{i+1} - t_i) = \mathcal{S}(g, \mathbb{P})$ □

Lemma 12.422. *Let $[a,b] \subseteq \mathbb{R}$, $\langle X, \|\|\rangle$ a normed real vector space and $f: [a,b] \rightarrow X$ a function, $\mathbb{P}_1 = (\{t_i^1\}_{i \in \{1,\ldots,n_1\}}, \{s_i^1\}_{i,\{1,\ldots,n_1-1\}})$, $\mathbb{P}_2 = (\{t_i^2\}_{i \in \{1,\ldots,n_2\}}, \{s_i^2\}_{i,\{1,\ldots,n_2-1\}})$ tagged partitions of $[a,b]$ then if $\mathcal{P} = \{t_i\}_{i \in \{1,\ldots,n\}} = \{t_i^1\}_{i \in \{1,\ldots,n_1\}} \boxplus \{t_i^2\}_{i \in \{1,\ldots,n_2\}}$ we have*

1. $\boldsymbol{S}(f, \mathbb{P}_1) = \sum_{i=1}^{n-1} f(s_{i_1(i)}^1) \cdot (t_{i+1} - t_i)$
2. $S(f, \mathbb{P}_2) = \sum_{i=1}^{n-1} f(s_{i_2(i)}^2) \cdot (t_{i+1} - t_i)$
3. $\mathcal{S}(f, \mathbb{P}_1) - \mathcal{S}(f, \mathbb{P}_2) = \sum_{i \in \{1,\ldots,n\}} (f(s_{i_1(i)}^1) - f(s_{i_2(i)}^2))$

(i_1, i_2 as defined in 12.417)

Proof. First we prove that for $j = 1, 2$ we have $\mathcal{S}(f, \mathbb{P}_j) = \sum_{i=1}^{n-1} f(s^j{}_{i_j(i)})(t_{i+1} - t_i)$. First as by 12.417 we have that i_j is surjective so that $\{1,\ldots,n-1\} = \bigcup_{k \in \{1,\ldots,n_j-1\}} i_j^{-1}(\{k\})$ and $\forall k, l \in \{1,\ldots,n_j-1\}$ with $k \neq j$ we have $i_j^{-1}(\{k\}) \bigcap i_j^{-1}(\{l\}) = \emptyset$ and $i^{-1}{}_j(\{k\}) = \{m_k^j, \ldots, M_k^j\}$ and $t_{m_k^j}^j = t_k$ and $t_{M_k^j+1}^j = t_{k+1}$, so if $i \in \{m_k^j, \ldots, M_k^j\}$ then $k = i_j(i)$. Using 10.47 we have $\sum_{i=1}^n f(s_{i_j(i)}^j) \cdot (t_{i+1} - t_i) = \sum_{i \in \{1,\ldots,n\}} f(s_{i_j(i)}^j) \cdot (t_{i+1} - t_i) = \sum_{k \in \{1,\ldots,n_j-1\}} \left(\sum_{i \in \{m_k^j,\ldots,M_k^j\}} f(s_{i_j(i)}^j) \cdot (t_{i+1} - t_i)\right) = \sum_{k \in \{1,\ldots,n_j-1\}} \left(\sum_{i \in \{m_k^j,\ldots,M_k^j\}} f(s_k^j) \cdot (t_{i+1} - t_i)\right) \underset{10.123}{=} \sum_{k \in \{1,\ldots,n_j-1\}} f(s_k^j) \cdot \left(\sum_{i=m_k^j}^{M_k^j} (t_{i+1} - t_i)\right) \underset{10.27}{=} \sum_{k \in \{1,\ldots,n_j-1\}} f(s_k^j) \cdot (t_{M_k^j+1} - t_{m_k^j}) = \sum_{k=1}^{n_j-1} f(s_k^j) \cdot (t_{k+1}^j - t_k^j) = \mathcal{S}(f, \mathbb{P}_j)$.

Finally $\mathcal{S}(f, \mathbb{P}_1) - \mathcal{S}(f, \mathbb{P}_2) = \left(\sum_{i=1}^n f(s^1{}_{i_1(i)})(t_{i+1} - t_i)\right) - \left(\sum_{i=1}^n f(s^2{}_{i_2(i)})(t_{i+1} - t_i)\right) = \sum_{i=1}^n (f(s_{i_1(i)}^1) - f(s_{i_2(i)}^2)) \cdot (t_{i+1} - t_i)$ □

Lemma 12.423. *Let $[a,b] \subseteq \mathbb{R}$, $\langle X, \|\|\rangle$ a normed real vector space and $f: [a,b] \rightarrow X$ a continuous function then for every $\varepsilon > 0$ there exists a $\delta < 0$ such that $\|\mathcal{S}(f, \mathbb{P}_1) - \mathcal{S}(f, \mathbb{P}_2)\| < \varepsilon$ for all $\mathbb{P}_1, \mathbb{P}_2$ such that $\mu(\mathbb{P}_1) < \delta$, $\mu(\mathbb{P}_2) < \delta$*

Proof. Using the fact that $[a,b]$ is compact in $\langle \mathbb{R}, || \rangle$ (see 12.257), f is continuous and 12.258 we have that f is uniform continuous on $[a,b]$. So there exists a $\delta > 0$ such that $\forall x, x' \in [a,b]$ with $|x - x'| < 2 \cdot \delta$ we have $\|f(x) - f(x')\| < \frac{\varepsilon}{b-a}$. Take now $\mathbb{P}_1, \mathbb{P}_2$ such that $\mu(\mathbb{P}_1) < \delta$, $\mu(\mathbb{P}_2) < \delta$ then by using the previous lemma we have

$$\|\mathcal{S}(f, \mathbb{P}_1) - \mathcal{S}(f, \mathbb{P}_2)\| = \left\| \sum_{i=1}^{n-1} (f(s^1_{i_1(i)}) - f(s^2_{i_2(i)})) \cdot (t_{i+1} - t_i) \right\| \tag{12.96}$$

As $t_i \in [t_i, t_{i+1}] \subseteq [t^1_{i_1(i)}, t^1_{i_1(i)+1}] \bigcap [t^2_{i_2(i)}, t^2_{i_2(i)+1}]$ we have as also $s^1_{i_1(i)} \in [t^1_{i_1(i)}, t^1_{i_1(i)+1}]$, $s^2_{i_2(i)} \in [t^2_{i_2(i)}, t^2_{i_2(i)+1}]$ that $|s^1_{i_1(i)} - s^2_{i_2(i)}| \leqslant |s^1_{i_1(i)} - t_i| + |t_i - s^2_{i_2(i)}| \leqslant \mu(\mathbb{P}_1) + \mu(\mathbb{P}_2) < \delta + \delta = 2 \cdot \delta$ so that

$$\|f(s^1_{i_1(i)}) - f(s^2_{i_2(i)})\| < \frac{\varepsilon}{b-a} \tag{12.97}$$

Using 12.96 and 12.97 we have $\|\mathcal{S}(f, \mathbb{P}_1) - \mathcal{S}(f, \mathbb{P}_2)\| \leqslant \sum_{i=1}^{n-1} \|f(s^1_{i_1(i)}) - f(s^2_{i_2(i)})\| |t_{i+1} - t_1| \leqslant \frac{\varepsilon}{b-a} \cdot \sum_{i=1}^{n-1} |t_{i+1} - t_i| \underset{t_{i+1} > t_i}{=} \frac{\varepsilon}{b-a} \cdot \sum_{i=1}^{n-1} (t_{i+1} - t_i) \underset{10.27}{=} \frac{\varepsilon}{b-a} \cdot (t_n - t_1) = \frac{\varepsilon}{b-a} \cdot (b-a) = \varepsilon$ what we set out to prove. □

Lemma 12.424. *Let $[a,b] \subseteq \mathbb{R}$, $a < b$, $\langle X, \|\| \rangle$ a real normed space and $f: [a,b] \rightarrow X$ a function then if a $I \in X$ satisfies the following condition*

1. *$\forall \varepsilon > 0$ there exists a $\delta > 0$ such that for every Riemann $\mathcal{S}(f, \mathbb{P})$ with $\mu(\mathbb{P}) < \delta$ (where $\mathbb{P}$ is a tagged partition of $[a,b]$) we have $\|I - \mathcal{S}(f, \mathbb{P})\| < \varepsilon$*

then I is unique.

Proof. Suppose that I, I' satisfies (1) and $I \neq I'$ then $\varepsilon = \|I - I'\| > 0$ then by (1) there exists a $\delta > 0$ such that if $\mu(\mathbb{P}) < \delta$ then $\|I - \mathcal{S}(f, \mathbb{P})\| < \frac{\varepsilon}{2}$, $\|I' - \mathcal{S}(f, \mathbb{P})\| < \frac{\varepsilon}{2}$. This means $\varepsilon = \|I - I'\| \leqslant \|I - \mathcal{S}(f, \mathbb{P})\| + \|\mathcal{S}(f, \mathbb{P}) - I'\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \Rightarrow \varepsilon < \varepsilon$ a contradiction. The only conclusion left is that $I = I'$. □

Definition 12.425. *Let $[a,b] \subseteq \mathbb{R}$, $\langle X, \|\| \rangle$ and $f: [a,b] \rightarrow X$ a function so that there exists a $I(f) \in X$ with $\forall \varepsilon > 0$ there exists a $\delta > 0$ such that for every Riemann sum $\mathcal{S}(f, \mathbb{P})$ with $\mu(\mathbb{P}) < \delta$ (where $\mathbb{P}$ is a tagged partition of $[a,b]$) we have $\|I(f) - \mathcal{S}(f, \mathbb{P})\| < \varepsilon$. f is then called **Riemann Integrable** and $I(f)$ is called the **Riemann Integral of f** and noted by $\int_a^b f$. The set of integrable functions on $[a,b]$ is noted as $\mathcal{L}([a,b], X)$. The integral $\int$ is then a function $\int_a^b: \mathcal{L}([a,b], X) \rightarrow X$ defined by $f \rightarrow \int_a^b f$*

Note 12.426. Sometimes we note a function $f: [a,b] \rightarrow Y$ defined by $x \rightarrow f(x)$ where $f(x)$ is in general a expression depending on x, in this case we can note $\int_a^b f$ by $\int_a^b f(x)\mathrm{dx}$, this notation is especially useful if we have for example a function $f: [a,b] \times X \rightarrow Y$ defined by $(t,x) \rightarrow f(t,x)$ then if we define for a $x \in X$ g_x: $[a,b] \rightarrow Y$ by $t \rightarrow g_x(t) = f(t,x)$ then we can write $\int_a^b g_x$ as $\int_a^b f(t,x)\mathrm{dt}$ which avoids the extra definition of g_x. Another benefit is that if we have $f: [a,b] \rightarrow \mathbb{K}$ with $x \rightarrow f(x) = \frac{\cos(x) + 1}{x + 1}$ then again we can write for $\int_a^b f$ the expression $\int_a^b \frac{\cos(x) + 1}{x + 1}\mathrm{dx}$. On the other hand we have introduced a extra indeterminisme as $\int_a^b f(x)\mathrm{dx}$ is the same as $\int_a^b f(y)\mathrm{dy}$ ... , this is the reason that we avoid this notation for the rest of this chapter. However if needed we can fall back on this notation if we feel a need to it.

Theorem 12.427. *Let $[a, b] \subseteq \mathbb{R}$, $\langle X, \|\|\rangle$ a normed real space then we have that there exists a family of tagged partitions on $[a, b]$ $\{\mathbb{P}_i\}_{i\in\mathbb{N}}$ such that $\lim_{i\rightarrow\infty} \mu(\mathbb{P}_i) = 0$*

Proof. Let $k \in \mathbb{N}$ define then $\left\{t_i^{(k)}\right\}_{i\in\{1,\ldots,k+1\}}$ by $t_i^{(k)} = (i-1)\cdot\frac{b-a}{k} + a$ then we have $t_1^{(k)} = (1-1)\cdot\frac{b-a}{k} + a = a$ and $t_{k+1}^{(k)} = (k+1-1)\cdot\frac{b-a}{k} + a = \frac{k\cdot(b-a)+a\cdot k}{k} = \frac{k\cdot b - a\cdot k + a\cdot k}{k} = b$, further $t_{i+1}^{(k)} - t_i^{(k)} = ((i+1)-1)\cdot\frac{b-a}{k} + a - (i-1)\cdot\frac{b-a}{k} - a = i\cdot\frac{b-a}{k} - i\cdot\frac{b-a}{k} + \frac{b-a}{k} = \frac{b-a}{k} > 0$, so that using 12.319 we have $\lim_{k\rightarrow\infty}\left(t_{i+1}^{(k)} - t_i^{(k)}\right) = 0$. Define now $\mathbb{P}_k = \left\{\left\{t_i^{(k)}\right\}_{i\in\{1,\ldots,k+1\}}, \left\{t_i^{(k)}\right\}_{i\in\{1,\ldots,k\}}\right\}$ then we have $\mu(\mathbb{P}_k) = \frac{b-a}{k}$ and thus for $\{\mathbb{P}_k\}_{k\in\mathbb{N}}$ we have $\lim_{i\rightarrow\infty}\mu(\mathbb{P}_k) = 0$

□

Theorem 12.428. *Let $[a, b] \subseteq \mathbb{R}$, X, $f: [a, b] \rightarrow X$ a* ***continuous*** *function then the following are equivalent*

1. *f is Riemann integrable with integral $\int_a^b f$*
2. *There exists a $I \in X$ such that for every family of tagged partitions $\{\mathbb{P}_i\}_{i\in\mathbb{N}}$ with $\lim_{i\rightarrow\infty}\mu(\mathbb{P}_i) = 0$ we have $\lim_{i\rightarrow\infty}\mathcal{S}(f, \mathbb{P}_i) = I$. We have then that $I = \int_a^b f$*

Proof.

1. **(⇒)** Let f be Riemann integrable with integral $\int_a^b f$ and take $\{\mathbb{P}_i\}_{i\in\mathbb{N}}$ such that $\lim_{i\rightarrow\infty}\mu(\mathbb{P}_i) = 0$. Then if $\varepsilon > 0$ there exists a $\delta > 0$ such that if $\mathbb{P}$ is such that $\mu(\mathbb{P}) < \delta$ then $\left\|\mathcal{S}(f, \mathbb{P}) - \int_a^b f\right\| < \varepsilon$. Now as $\lim_{i\rightarrow\infty}\mu(\mathbb{P}_i) = 0$ then there exists a $N \in \mathbb{N}$ such that if $n \geqslant N$ then $\mu(P_n) = |\mu(\mathbb{P}_n) - 0| < \delta \Rightarrow \big\|\mathcal{S}(f, \mathbb{P}_i) - \int_a^b f\big\| < \varepsilon$ proving that $\lim_{i\rightarrow\infty}\mathcal{S}(f, \mathbb{P}_i) = \int_a^b f$

2. **(⇐)** First by 12.427 there exists a $\{\mathbb{P}_i\}_{i\in\mathbb{N}}$ with $\lim_{i\rightarrow\infty}\mu(\mathbb{P}_i) = 0$, by the hypothesis we have then that $I = \lim_{i\rightarrow\infty}\mathcal{S}(f, \mathbb{P}_i)$ exists. Take now $\varepsilon > 0$ then there exists a $N_1 \in \mathbb{N}$ such that if $n \geqslant N_1$ then $\|\mathcal{S}(f, \mathbb{P}_i) - I\| < \frac{\varepsilon}{2}$, using 12.423 there exists a $\delta > 0$ such that if $\mathbb{P}', \mathbb{P}''$ are two tagged partitions with $\mu(\mathbb{P}') < \delta$, $\mu(\mathbb{P}'') < \delta$ then $\|\mathcal{S}(f, \mathbb{P}') - \mathcal{S}(f, \mathbb{P}'')\| < \frac{\varepsilon}{2}$, as $\lim_{i\rightarrow\infty}\mu(\mathbb{P}_i) = 0$ there exists also a N_2 so that if $n \geqslant N_2$ then $\mu(\mathbb{P}_n) = |\mu(\mathbb{P}_n) - 0| < \delta$. Take now $N = \max(N_1, N_2) \geqslant N_1, N_2$ then $\mu(\mathbb{P}_N) < \delta$ and $\|\mathcal{S}(f, \mathbb{P}_N) - I\| < \frac{\varepsilon}{2}$ and assume that for a partition $\mathbb{P}$ we have $\mu(\mathbb{P}) < \delta$ then $\|\mathcal{S}(f, \mathbb{P}) - I\| = \|\mathcal{S}(f, \mathbb{P}) - \mathcal{S}(f, \mathbb{P}_N)\| + \|\mathcal{S}(f, \mathbb{P}_n) - I\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. Proving that f is integrable and as we have uniqueness by 12.424 we must have $I = \int_a^b f$. □

Theorem 12.429. *Let $f, g \in \mathcal{L}([a, b], \mathbb{R})$ such that $\forall x \in [a, b]$ we have $f(x) \leqslant g(x)$ then $\int_a^b f \leqslant \int_a^b g$ then $\int_a^b f \leqslant \int_a^b g$*

Proof. Using 12.427 there exists a $\{\mathbb{P}_i\}_{i\in\mathbb{N}}$ such that $\lim_{i\rightarrow\infty}\mu(\mathbb{P}_i)=0$ then we have

$$\begin{aligned}\int_a^b f &\underset{12.428}{=} \lim_{i\rightarrow\infty}\mathcal{S}(f,\mathbb{P}_i)\\ &\leqslant_{12.343 \text{ and } 12.421} \lim_{i\rightarrow\infty}\mathcal{S}(g,\mathbb{P}_i)\\ &\underset{12.428}{=} \int_a^b g\end{aligned}$$

□

We prove now that $\mathcal{L}([a,b],X)$ is not empty if X is a Banach space and f is continuous.

Theorem 12.430. *Let $[a,b]\subseteq\mathbb{R}$, $\langle X,\|\|\rangle$ a real normed Banach space and f: $[a,b]\rightarrow X$ a continuous function then f is integrable. In other words we have that $\mathcal{C}([a,b],X)\subseteq\mathcal{L}([a,b],X)$*

Proof. Let $\{\mathbb{P}_i\}_{i\in\mathbb{N}}$ be a family of tagged partitions on $[a,b]$ with $\lim_{i\rightarrow\infty}\mu(\mathbb{P}_i)=0$. Given a $\varepsilon>0$ we can find by 12.423 there exists a $\delta>0$ such that if $\mu(\mathbb{P})<\delta$, $\mu(\mathbb{P}')<\delta$ then $\|\mathcal{S}(f,\mathbb{P})-\mathcal{S}(f,\mathbb{P}')\|<\varepsilon$, we can then find a $N\in\mathbb{N}$ such that if $i\geqslant N$ we have $\mu(\mathbb{P}_i)=|\mu(\mathbb{P}_i)-0|<\delta$. So if $n,m\geqslant\mathbb{N}_0$ then $\mu(\mathbb{P}_n),\mu(\mathbb{P}_m)<\delta$ and thus $\|\mathcal{S}(f,\mathbb{P}_n)-\mathcal{S}(f,\mathbb{P}_m)\|<\varepsilon$ proving that $\{\mathcal{S}(f,\mathbb{P}_i)\}_{i\in\mathbb{N}}$ is a Cauchy sequence in X. As X is a Banach space we have that $\{\mathcal{S}(f,\mathbb{P}_i)\}_{i\in\mathbb{N}}$ converges to a limit I. Using then the previous theorem (see 12.428) we have then that f is Integrable with $\lim_{i\rightarrow\infty}\mathcal{S}(f,\mathbb{P}_i)=\int_a^b f$ □

Up to now we have always assumed that $a<b$ let's now remove this restriction

Definition 12.431. *Let $[a,b]\subseteq\mathbb{R}$ with $a\leqslant b$, $\langle X,\|\|\rangle$ a real normed space then f: $[a,b]\rightarrow X$ is Riemann Integrable with integral $\int_a^b$ if and only if*

1. *If $a<b$ then we use the definition of 12.425*
2. *If $a=b$ then f is always integrable and $\int_a^b f=0$*

From now on if we say $[a,b]\in\mathbb{R}$ then $a\leqslant b$ (so $a=b$ is back allowed)

Theorem 12.432. *Let $[a,b]\subseteq\mathbb{R},a\leqslant b$ and $\langle X,\|\|\rangle$ a real normed space then $\int_a^b$: $\mathcal{L}([a,b],X)\rightarrow X$ is linear. In other words $\forall\alpha,\beta\in\mathbb{R}$ and $\forall f,g\in\mathcal{L}([a,b],X)$ we have $\alpha\cdot f+\beta\cdot g\in\mathcal{L}([a,b],X)$ and $\int_a^b(\alpha\cdot f+\beta\cdot g)=\alpha\cdot\int_a^b f+\beta\cdot\int_a^b g$*

Proof. Consider the two cases for the integral

1. **($\boldsymbol{a=b}$)** Then the integral exists always and $\int_a^b(\alpha\cdot f+\beta\cdot g)=0+0=\alpha\cdot 0+\beta\cdot 0=\alpha\cdot\int_a^b f+\beta\cdot\int_a^b g$

2. **$(\boldsymbol{a<b})$** Let $\mathbb{P}=(\{t_i\}_{i\in\{1,\ldots,n\}},\{s_i\}_{i\in\{1,\ldots,n-1\}})$ a tagged partition of $[a,b]$ then $\mathcal{S}(\alpha\cdot f+\beta\cdot g,\mathbb{P})=\sum_{i=1}^{n}(\alpha\cdot f+\beta\cdot g)(s_i)\cdot(t_{i+1}-t_i)=\sum_{i=1}^{n}(\alpha\cdot f(s_i)+\beta\cdot g(s_i))\cdot(t_{i+1}-t_i)=\alpha\cdot\sum_{i=1}^{n}f\cdot(t_{i+1}-t_i)+\beta\cdot\sum_{i=1}^{n}g\cdot(t_{i+1}-t_i)=\alpha\cdot\mathcal{S}(f,\mathbb{P})+\beta\cdot\mathcal{S}(g,\mathbb{P})$ giving

$$\mathcal{S}(\alpha\cdot f+\beta\cdot g,\mathbb{P})=\alpha\cdot\mathcal{S}(f,\mathbb{P})+\beta\cdot\mathcal{S}(g,\mathbb{P})\tag{12.98}$$

Let $\varepsilon>0$ then as f,g is integrable there exists a $\delta>0$ such that if $\mu(\mathbb{P})<\delta$ then $\left\|\int_a^b f-\mathcal{S}(f,\mathbb{P})\right\|<\frac{\varepsilon}{2\cdot(|\alpha|+1)}$ and $\left\|\int_a^b g-\mathcal{S}(g,\mathbb{P})\right\|<\frac{\varepsilon}{2\cdot(|\beta|+1)}$ and thus $\left\|\left(\alpha\cdot\int_a^b f+\beta\cdot\int_a^b g\right)-\mathcal{S}(\alpha\cdot f+\beta\cdot g,\mathbb{P})\right\|\underset{12.98}{=}\left\|\alpha\cdot\int_a^b f+\beta\cdot\int_a^b g-\alpha\cdot\mathcal{S}(f,\mathbb{P})-\beta\cdot\mathcal{S}(g,\mathbb{P})\right\|\leqslant\left\|\alpha\cdot\left(\int_a^b f-\mathcal{S}(f,\mathbb{P})\right)\right\|+\left\|\beta\cdot\left(\int_a^b g-\mathcal{S}(g,\mathbb{P})\right)\right\|=|\alpha|\cdot\left\|\int_a^b f-\mathcal{S}(f,\mathbb{P})\right\|+|\beta|\cdot\left\|\int_a^b g-\mathcal{S}(g,\mathbb{P})\right\|<|\alpha|\cdot\frac{\varepsilon}{2\cdot(|\alpha|+1)}+|\beta|\cdot\frac{\varepsilon}{2\cdot(|\beta|+1)}<\varepsilon$ proving that $\int_a^b(\alpha\cdot f+\beta\cdot g)=\alpha\cdot\int_a^b f+\beta\cdot\int_a^b g$. □

Theorem 12.433. *Let $[a,b]\subseteq\mathbb{R}$, $a\leqslant b$ and $\langle X,\|\|\rangle$ a real normed space and $f\in\mathcal{L}([a,b],X)$ then we have*

1. *$\forall\varphi\in L(X,\mathbb{R})$ (a continuous linear function) we have that $\varphi\circ f$ is integrable and $\int_a^b(\varphi\circ f)=\varphi\left(\int_a^b f\right)$*
2. *If $\|f\|$ is continuous then $f,\|f\|$ are integrable and $\left\|\int_a^b f\right\|\leqslant\int_a^b\|f\|$*
3. *Given $c\in X$ the constant function c defined by $x\to c(x)=c$ is integrable and $\int_a^b c=c\cdot(b-a)$*
4. *If f is continuous and $m\in\mathbb{R}$ is such that $\forall x\in[a,b]$ we have $\|f(x)\|<m$ then $\left\|\int_a^b f\right\|\leqslant m\cdot(b-a)$*

Proof. We divide the proof in the cases of $a=b$ and $a<b$

1. **$(\boldsymbol{a=b})$**
   a. If $\varphi\in L(X,\mathbb{R})$ then we have $\int_a^a\varphi\circ f=0\underset{10.173}{=}\varphi(0)=\varphi(\int_a^a f)$
   b. $\|\int_a^a f\|=\|0\|=0\leqslant 0=\int_a^a\|f\|$
   c. $\int_a^a c=0=c\cdot 0=c\cdot(a-a)$
   d. $\|\int_a^a f\|=\|0\|\leqslant 0=m\cdot 0=m\cdot(a-a)$
2. **$(\boldsymbol{a<b})$**
   a. If $\mathbb{P}=(\{t_i\}_{i\in\{1,\ldots,n\}},\{s_i\}_{i\in\{1,\ldots,n-1\}})$ then we have $\mathcal{S}(\varphi\circ f,\mathbb{P})=\sum_{i=1}^{n-1}\varphi(f(s_i))\cdot(t_{i+1}-t_i)\underset{\varphi\,\text{is linear}}{=}\varphi\left(\sum_{i=1}^{n-1}f(s_i)\cdot(t_{i+1}-t_i)\right)=\varphi(\mathcal{S}(f,\mathbb{P}))$, so we have

$$\mathcal{S}(\varphi\circ f,\mathbb{P})=\varphi(\mathcal{S}(f,\mathbb{P}))\tag{12.99}$$

Now as φ is continuous, we have that given a $\varepsilon>0$ there exists a $\delta'>0$ such that if $\|x\|=\|x-0\|<\delta'$ then $|\varphi(x)|<\varepsilon$. As f is integrable there exists a $\delta>0$ such that if $\mu(\mathbb{P})<\delta$ then $\left\|\int_a^b - \mathcal{S}(f,\mathbb{P})\right\|<\delta'$ and thus $\left|\varphi\left(\int_a^b f\right)-\mathcal{S}(\varphi\circ f,\mathbb{P})\right| \underset{12.99}{=} \left|\varphi\left(\int_a^b f\right)-\varphi(\mathcal{S}(f,\mathbb{P}))\right| \underset{\varphi \text{ is linear}}{=} \left|\varphi\left(\left(\int_1^b f\right)-\mathcal{S}(f,\mathbb{P})\right)\right|<\varepsilon$ proving that $\varphi\circ f$ is integrable with integral $\varphi\left(\int_a^b f\right)$.

b. As f is continuous we have that f is integrable, further given $x\in X$ and $\varepsilon>0$ a $\delta>0$ so that if $\|x-y\|<\delta$ then $\|f(x)-f(y)\|<\varepsilon\Rightarrow |\|f(x)\|-\|f(y)\||\leqslant\|f(x)-f(y)\|<\varepsilon$ proving that $\|f\|$ is continuous and thus integrable. Second given $\mathbb{P}=(\{t_i\}_{i\in\{1,\ldots,n\}},\{s_i\}_{i\in\{1,\ldots,n-1\}})$ we have $\|\mathcal{S}(f,\mathbb{P})\|=\|\sum_{i=1}^{n-1} f(s_i)\cdot(t_{i+1}-t_i)\|\leqslant\sum_{i=1}^{n-1}\|f(s_i)\cdot(t_{i+1}-t_i)\|=\sum_{i=1}^{n-1}\|f(s_i)\|\cdot(t_{i+1}-t_i)$ proving that

$$\|\mathcal{S}(f,\mathbb{P})\|\leqslant\mathcal{S}(\|f\|,\mathbb{P}) \tag{12.100}$$

This means that if $\{\mathbb{P}_i\}_{i\in\mathbb{N}}$ is such that $\lim_{i\to\infty}\mu(\mathbb{P}_i)=0$ then $\left\|\int_a^b f\right\|=\left\|\lim_{i\to\infty}\mathcal{S}(f,\ \mathbb{P})\right\| \underset{\|\| \text{ is continuous and } 12.323}{=} \lim_{i\to\infty}\|\mathcal{S}(f,\mathbb{P})\| \leqslant_{12.100 \text{ and } 12.339} \lim_{i\to\infty}\mathcal{S}(\|f\|,\mathbb{P})=\int_a^b\|f\|$.

c. If c is the constant function then we have for $\mathbb{P}=(\{t_i\}_{i\in\{1,\ldots,n\}},\{s_i\}_{i\in\{1,\ldots,n-1\}})$ that $\mathcal{S}(c,\mathbb{P})=\sum_{i=1}^{n-1} c(s_i)\cdot(t_{i+1}-t_i)=\sum_{i=1}^{n-1} c\cdot(t_{i+1}-t_i)=c\cdot\sum_{i=1}^{n-1}(t_{i+1}-t_i) \underset{10.27}{=} c\cdot(t_n-t_1)=c\cdot(b-a)$

d. First $\mathbb{P}=(\{t_i\}_{i\in\{1,\ldots,n\}},\{s_i\}_{i\in\{1,\ldots,n-1\}})$ is a tagged partition on $[a,b]$ then we have $\|\mathcal{S}(f,\mathbb{P})\|=\|\sum_{i=1}^{n-1} f(s_i)\cdot(t_{i+1}-t_i)\|\leqslant\sum_{i=1}^{n-1}\|f(s_i)\|\cdot(t_{i+1}-t_i)\leqslant\sum_{i=1}^{n-1} m\cdot(t_{i+1}-t)=m\cdot\sum_{i=1}^{n-1}(t_{i+1}-t_i) \underset{10.27}{=} m\cdot(b-a)$ proving that

$$\|\mathcal{S}(f,\mathbb{P})\|\leqslant m\cdot(b-a) \tag{12.101}$$

If now $\{\mathbb{P}_i\}_{i\in\mathbb{N}}$ is such that $\lim_{i\to\infty}\mu(\mathbb{P}_i)=0$ then $\left\|\int_a^b f\right\|=\left\|\lim_{i\to\infty}\mathcal{S}(f,\mathbb{P}_i)\right\| \underset{\|\| \text{ is continuous and } 12.323}{=} \lim_{i\to\infty}\|\mathcal{S}(f,\mathbb{P}_i)\|\leqslant_{12.101 \text{ and } 12.337} m\cdot(b-a)$ □

Lemma 12.434. *Let $c\in[a,b]\subseteq\mathbb{R}$, $a<c<b$ and let $\mathbb{P}_1=(\{t_i^1\}_{i\in\{1,\ldots,n_1\}},\{s_i^2\}_{i\in\{1,\ldots,n_1-1\}})$ be a tagged partition on $[a,c]$, $\mathbb{P}_2=(\{t_i^2\}_{i\in\{1,\ldots,n_2\}},\{s_i^2\}_{i\in\{1,\ldots,n_2-1\}})$ a tagged partition on $[c,b]$ then $\mathbb{P}=\mathbb{P}_1\bigcup\mathbb{P}_2$ defined by $\mathbb{P}=(\{t_i\}_{i\in\{1,\ldots,n_1+n_2-1\}},\{s_i\}_{i\in\{1,\ldots,n_1+n_2-2\}})$ where $t_i=\begin{cases} t_i^1 \text{ if } i\in\{1,\ldots,n_1\} \\ t_{i-n_1+1}^2 \text{ if } i\in\{n_1+1,\ldots,n_1+n_2-1\}\end{cases}$ and $s_i=\begin{cases} s_i^1 \text{ if } i\in\{1,\ldots,n_1-1\} \\ s_{i-n_1+1}^2 \text{ if } i\in\{n_1,\ldots,n_1+n_2-2\}\end{cases}$ then $\mathbb{P}$ is a tagged partition of $[a,b]$ with $\mu(\mathbb{P})=\max(\mu_1(\mathbb{P}_1),\mu(\mathbb{P}_2))$. If $\langle X,\|\|\rangle$ is a normed space and $f:[a,b]\to X$ a function then $\mathcal{S}(f,\mathbb{P})=\mathcal{S}(f,\mathbb{P}_1)+\mathcal{S}(f,\mathbb{P}_2)$*

Proof. To prove this note:

1. $t_1=t_1^1=a$

2. $t_{n_1+n_2-1} \underset{1<n_2\Rightarrow 2\leqslant n_2\Rightarrow 2+(n_1-1)\leqslant n_2+(n_1-1)+n_2\Rightarrow n_1+1\leqslant n_1+n_2-1}{=} t^2_{(n_1+n_2-1)-n_1+1} = t^2_{n_2} = b$

3. If $i \in \{1, ..., n_1+n_2-2\}$ then we have either

   a. $\boldsymbol{(i \in \{1, ..., n_1-1\})}$ then $s_i = s^1_i \in [t^1_i, t^1_{i+1}] \underset{1\leqslant i, i+1\leqslant n_1}{=} [t_i, t_{i+1}]$

   b. $\boldsymbol{(i = n_1)}$ then $s_{n_1} = s^2_{(n_1-n_1+1)} = s^2_1 \in [t^2_1, t^2_{1+1}] = [c, t^2_1] = [t^1_{n_1}, t^2_2] = [t_{n_1}, t^2_{(n_1+1)-n_1+1}] = [t_{n_1}, t_{n_1+1}]$

   c. $\boldsymbol{(i \in \{n_1+1, ..., n_1+n_2-1\})}$ then $s_i = s^2_{i-n_1+1} \in [t^2_{i-n_1+1}, t^2_{(i-n_1+1)+1}] = [t_i, t_{i+1}]$

4. $\{|t_{i+1} - t_i| | i \in \{1, ..., n_1+n_2-2\}\} = \{|t_{i+1} - t_i| | i \in \{1, ..., n_1-1\}\} \bigcup \{|t_{i+1} - t_i| | i \in \{n_1\}\} \bigcup \{|t_{i+1} - t_i| | i \in \{n_1+1, ..., (n_1+n_2-1)-1\}\} = \{|t^1_{i+1} - t^1_i| | i \in \{1, ..., n_1-1\}\} \bigcup \{|t^2_{(n_1+1)-n_1+1} - t^1_{n_1}|\} \bigcup \{|t^2_{(i-n_1+1)+1} - t^2_{i-n_1+1}| | i \in \{n_1+1, ..., n_1+n_2-2\}\} = \{|t^1_{i+1} - t^1_i| | i \in \{1, ..., n_1-1\}\} \bigcup \{|t^2_2 - c|\} \bigcup \{|t^2_{((j+(n_1+1)-2)-n_1+1)+1} - t^2_{(j+(n_1+1)-2)-n_1+1}| | j \in \{2, ..., n_2-1\}\} = \{|t^1_{i+1} - t^1_i| | i \in \{1, ..., n_1-1\}\} \bigcup \{|t^2_2 - t^2_1|\} \bigcup \{|t^2_{j+1} - t^2_{(j}| | j \in \{2, ..., n_2-1\}\} = \{|t^1_{i+1} - t^1_i| | i \in \{1, ..., n_1-1\}\} \bigcup \{|t^2_{j+1} - t^2_{(j}| | j \in \{1, ..., n_2-1\}\}$ so that $\mu(\mathbb{P}) = \max(\{|t_{i+1} - t_i| | i \in \{1, ..., n_1+n_2-2\}\}) = \max(\{\max(\{|t^1_{i+1} - t^1_i| | i \in \{1, ..., n_1-1\}\}), \max(\{|t^2_{j+1} - t^2_{(j}| | j \in \{1, ..., n_2-1\}\})\}) = \max(\mu(\mathbb{P}_1), \mu(\mathbb{P}_2))$

5. $\mathcal{S}(f, \mathbb{P}) = \sum_{i=1}^{(n_1+n_2-1)-1} f(s_i) \cdot (t_{i+1} - t_i) = \left(\sum_{i=1}^{n_1-1} f(s_i) \cdot (t_{i+1} - t_i)\right) + \left(\sum_{i=n_1}^{n_1} f(s_i) \cdot (t_{i+1} - t_i)\right) + \left(\sum_{i=n_1+1}^{(n_1+n_2-1)-1}\right) = \left(\sum_{i=1}^{n_1-1} f(s^1_i) \cdot (t^1_{i+1} - t^1_i)\right) + (f(s_{n_1}) \cdot (t_{n_1+1} - t_{n_1})) + \left(\sum_{i=n_1+1}^{(n_1+n_2-1)-1} f(s^2_{i-n_1+1}) \cdot (t^2_{(i-n_1+1)+1} - t^2_{(i-n_1+1)})\right) = \left(\sum_{i=1}^{n_1-1} f(s^1_i) \cdot (t^1_{i+1} - t^1_i)\right) + (f(s^2_1) \cdot (t^2_{(n_1-n_1+1)+1} - t^1_{n_1})) + \left(\sum_{j=2}^{n_2-1} f(s^2_{j+(n_1+1)-2-n_1+1}) \cdot (t^2_{((j+(n_1+1)-2)-n_1+1)+1} - t^2_{(j+(n_1+1)-n_1+1)-2})\right) = \left(\sum_{i=1}^{n_1-1} f(s^1_i) \cdot (t^1_{i+1} - t^1_i)\right) + (f(s_{n_1}) \cdot (t^2_2 - c)) + \left(\sum_{j=2}^{n_2-1} f(s^2_j) \cdot (t^2_{j+1} - t^2_j)\right) = \left(\sum_{i=1}^{n_1-1} f(s^1_i) \cdot (t^1_{i+1} - t^1_i)\right) + (f(s_{n_1}) \cdot (t^2_2 - t^2_1)) + \left(\sum_{j=2}^{n_2-1} f(s^2_j) \cdot (t^2_{j+1} - t^2_j)\right) = \left(\sum_{i=1}^{n_1-1} f(s^1_i) \cdot (t^1_{i+1} - t^1_i)\right) + \left(\sum_{j=1}^{n_2-1} f(s^2_j) \cdot (t^2_{j+1} - t^2_j)\right) = \mathcal{S}(f, \mathbb{P}_1) + \mathcal{S}(f, \mathbb{P}_2)$ □

Theorem 12.435. *Let $[a, b] \subseteq \mathbb{R}, a \leqslant b, \langle X, \|\|\rangle$ a real metric space and $c \in [a, b]$ and $f \in \mathcal{C}([a, b], X)$ then $f \in \mathcal{C}([a, c], X)$ and $f \in \mathcal{C}([c, d], X)$ so that $f: [a, b] \rightarrow X$, $f: [a, c] \rightarrow X$ and $f: [c, b] \rightarrow X$ are integrable and $\int_a^c f + \int_c^b f = \int_a^b f$*

Proof. We have the following cases to consider:

1. $\boldsymbol{(a = b)}$ Then $a = c = b$ so that $\int_a^b f = \int_a^a = 0 = 0 + 0 = \int_a^a f + \int_a^a f = \int_a^c f + \int_c^b$

2. $\boldsymbol{(a < b \wedge a = c)}$ Then $\int_a^b f = 0 + \int_a^b = \int_a^a f + \int_a^b f = \int_a^c f + \int_c^b f$

3. $\boldsymbol{(a < b \wedge c = b)}$ Then $\int_a^b f = \int_a^b f + 0 = \int_a^b f + \int_b^b = \int_a^c f + \int_c^b$

4. **($\boldsymbol{a<c\wedge c<b}$)** As $f:[a,c]\to X$ and $f:[c,b]\to X$ are continuous we have that they are integrable so that by 12.428 there exists a $\{\mathbb{P}_i^1\}_{i\in\mathbb{N}}$ (tagged partitions on $[a,c]$) and $\{\mathbb{P}_i^2\}_{i\in\mathbb{N}}$ (tagged partitions on $[c,b]$) with $\lim_{i\to\infty}\mu(\mathbb{P}_i^1)=0$, $\lim_{i\to\infty}\mu(\mathbb{P}_i^2)=0$ so that $\lim_{i\to\infty}\mathcal{S}(f,\mathbb{P}_i^1)=\int_a^c f$ and $\lim_{i\to\infty}\mathcal{S}(f,\mathbb{P}_i^2)=\int_c^b f$. Take then $\{\mathbb{P}_i^1\bigcup\mathbb{P}_i^1\}_{i\in\mathbb{N}}$ then by 12.434 we have that $\{\mathbb{P}_i^1\bigcup\mathbb{P}_i^2\}_{i\in\{1,\ldots,n\}}$ is a family of tagged partitions of $[a,b]$. Also if $\varepsilon>0$ then $\exists N_1,N_2$ such that if $i\geqslant N_1$ then $\mu(\mathbb{P}_i^1)<\frac{\varepsilon}{2}$, if $i\geqslant N_2$ then $\mu(\mathbb{P}_i^2)<\frac{\varepsilon}{2}$. So if $i\geqslant N=\max(N_1,N_2)$ then $|\mu(\mathbb{P}_i^1\bigcup\mathbb{P}_i^i)-0|=|\mu(\mathbb{P}_i^1\bigcup\mathbb{P}_i^2)|\underset{12.434}{=}|\mu(\mathbb{P}_i^1)+\mu(\mathbb{P}_i^2)|\leqslant|\mu(\mathbb{P}_i^1)|+|\mu(\mathbb{P}_i^2)|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$ proving that $\lim_{i\to\infty}\mu(\mathbb{P}_i^1\bigcup\mathbb{P}_i^2)=0$. Next $\int_a^c f+\int_c^b f=\lim_{i\to\infty}\mathcal{S}(f,\mathbb{P}_i^1)+\lim_{i\to\infty}\mathcal{S}(f,\mathbb{P}_i^2)\underset{12.341}{=}\lim_{i\to\infty}(\mathcal{S}(f,\mathbb{P}_i^1)+\mathcal{S}(f,\mathbb{P}_i^2))\underset{12.434}{=}\lim_{i\to\infty}\mathcal{S}(f,\mathbb{P}_i^1\bigcup\mathbb{P}_i^2)\underset{12.434}{=}\int_a^b f$

$\square$

Theorem 12.436. *Let $[a,b]\subseteq\mathbb{R}, a\leqslant b$, $\langle X,\|\|_X\rangle$ a real normed space, $\langle Y,\|\|_Y\rangle$ a real Banach space (so that by 12.352 $L(X,Y)$ is a real Banach space, $f\in\mathcal{C}([a,b],L(X,Y))$. If we define $\forall x\in X$ $f(.)(x):[a,b]\to Y$ by $f(.)(x)(t)=f(t)(x)$ then we have $\left(\int_a^b f\right)(x)=\int_a^b f(.)(x)$*

Proof. First note that given a $x\in X$ the function $\varphi_x:L(X,Y)\to Y$ defined by $\varphi_x(f)=f(x)$ is continuous as $\|\varphi_x(f)-\varphi_x(g)\|_Y=\|f(x)-g(x)\|_X=\|(f-g)(x)\|\leqslant\|f-g\|\cdot\|x\|_X$ so if $\varepsilon>0$ take then $\delta=\frac{\varepsilon}{\|x\|_X+1}$ so that if $\|f-g\|<\delta$ then $\|\varphi_x(f)-\varphi_x(g)\|<\frac{\|x\|_X\cdot\varepsilon}{\|x\|_X+1}<\varepsilon$. Second φ_x is linear as is proved by $\varphi_x(\alpha\cdot f+\beta\cdot g)=(\alpha\cdot f+\beta\cdot g)(x)=\alpha\cdot f(x)+\beta\cdot g(x)=\alpha\cdot\varphi_x(f)+\beta\cdot\varphi_x(g)$. Note that if $f\in\mathcal{C}([a,b],L(X,Y))$ we have that $\forall t\in[a,b]$ that $f(.)(x)(t)=f(t)(x)=\varphi_x(f(t))=(\varphi_x\circ f)(t)$ so that $f(.)(x)=\varphi_x\circ f:[a,b]\to Y$ is continuous as φ_x and f is continuous. From the continuity it follows that $f(.)(x)$ is integrable and thus that $\int_a^b f(.)(x)$ exists. Thirdly if $\mathbb{P}=(\{t_i\}_{i\in\{1,\ldots,n\}},\{s_i\}_{i\in\{1,\ldots,n-1\}})$ is a tagged partition of $[a,b]$ then $\mathcal{S}(f(.)(x),\mathbb{P})=\mathcal{S}(\varphi_x\circ f,\mathbb{P})=\sum_{i=1}^n(\varphi_x\circ f)(s_i)\cdot(t_{i+1}-t_i)=\sum_{i=1}^n\varphi_x(f(s_i))\cdot(t_{i+1}-t_i)=\varphi_x(\sum_{i=1}^n f(s_i)\cdot(t_{i+1}-t_i))=\varphi_x(\mathcal{S}(f,\mathbb{P}))$. So if $\{\mathbb{P}_i\}_{i\in\mathbb{N}}$ is a family of tagged partitions with $\lim_{i\to\infty}\mu(\mathbb{P}_i)=0$ then $\int_a^b f(.)(x)=\lim_{i\to\infty}\mathcal{S}(f(.)(x),\mathbb{P}_i)=\lim_{i\to\infty}\varphi_x(\mathcal{S}(f,\mathbb{P}_i))\underset{12.323}{=}\varphi_x\left(\lim_{i\to\infty}\mathcal{S}(f,\mathbb{P}_i)\right)=\varphi_x\left(\int_a^b f\right)=\left(\int_a^b f\right)(x)$ $\square$

12.13 Connected Sets

Definition 12.437. *A topological space $\langle X,\mathcal{T}\rangle$ is **connected** if $\forall U_1,U_2\in\mathcal{T}$ with $U_1\bigcap U_2=\emptyset$ and $X=U_1\bigcup U_2$ we have that either $U_1=\emptyset$ or $U_2=\emptyset$ (in other words X is not the union of two non empty disjoint open sets). The topological space is **disconnected** if it is not connected, or in other words there exists open $U_1,U_2\in\mathcal{T}$ with $U_1\neq\emptyset\neq U_2$ and $X=U_1\bigcup U_2$*

Theorem 12.438. *Let $\langle X, \mathcal{T}\rangle$ be a topological space then the following are equivalent*

1. *$\emptyset$ an X are the only subsets of X that are open and closed*
2. *X is connected*
3. *$\forall A_1, A_2$ closed sets with $A_1 \bigcap A_2 = \emptyset$ and $X = A_1 \bigcup A_2$ we have either $A_1 = \emptyset$ or $A_2 = \emptyset$*

Proof. We make in this proof heavy use of 1.32!

1. **(1 ⇒ 2)** Assume that U_1, U_2 are disjoint open sets with $X = U_1 \bigcup U_2$ then $U_2 = X \setminus U_1$ which is closed. According to (1) we must then have that either $U_2 = \emptyset$ or $U_2 = X$ and then $U_1 = X \setminus U_2 = X \setminus X = \emptyset$ so $U_1 = \emptyset$ proving that X is connected.
2. **(2 ⇒ 3)** Assume that we have two disjoint closed sets A_1, A_2 with $X = A_1 \bigcup A_2$ then $A_1 = X \setminus A_2$ is open and $A_2 = X \setminus A_1$ is open, so by the connectness of X we must have either $A_1 = \emptyset$ or $A_2 = \emptyset$.
3. **(3 ⇒ 1)** Let $A \subseteq X$ and A is open and closed then $X \setminus A$ is closed, $A \bigcap (X \setminus A) = \emptyset$ and $X = A \bigcup (X \setminus A)$. By (3) we must have either $A = \emptyset$ or $X \setminus A = \emptyset \Rightarrow A = X$. □

Definition 12.439. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and let $A \subseteq X$ then A is connected if it is connected in the subspace topology*

Theorem 12.440. *Let $\langle \mathbb{R}, ||\rangle$ be the real normed space then $\forall a, b \in \mathbb{R}$ with $a \leqslant b$ we have that $[a, b]$ is a connected subset.*

Proof. First as $[a, b] = \left[\frac{a+b}{2} - \frac{b-a}{2}, \frac{a+b}{2} + \frac{b-a}{2}\right] = \bar{B}_{||}\left(\frac{a+b}{2}, \frac{b-a}{2}\right)$ which is closed by 12.56 so that $[a, b]$ is a closed subset of $\mathbb{R}$. So if $A \subseteq [a, b]$ is closed in the subspace topology of $[a, b]$ then $A = A' \bigcap [a, b]$ with A' is closed in $\mathbb{R}$ so that A is also closed in $\mathbb{R}$. Assume now that $[a, b]$ is not connected then by the previous theorem we can find disjoint closed sets (closed in $[a, b]$ and thus also in $\mathbb{R}$) A_1, A_2 with $A_1, A_2 \neq \emptyset$ such that $[a, b] = A_1 \bigcup A_2$. We can always assume that $b \in A_2$ (otherwise switch $1 \leftrightarrow 2$). As $A_1 \subseteq A_1 \bigcup A_2 = [a, b]$ we have that A_1 is bounded above by b and is not empty so by order conditional completeness of the reals (see 9.43) $c = \sup(A_1)$ exists. If now there exists a $x, y \in \mathbb{R}$ with $c \in]x, y[$ then $x < c < y$ and by the definition of a supremum there exists a $a' \in A_1$ such that $x < a' \leqslant c < y \Rightarrow a' \in]x, y[\Rightarrow]x, y[\bigcap A_1 \neq \emptyset \underset{A_1 \text{ is closed}}{\Rightarrow} c \in A_1$. Since b is a upper bound of A_1, $b \in A_2$ and $A_1 \bigcap A_2 = \emptyset \Rightarrow b \notin A_1$ we must have $c < b \Rightarrow b - c > 0$. Now if $x \in]c, b]$ then $x > c$ and thus $x \notin A_1$ $(c = \sup(A_1)) \Rightarrow]c, b] \subseteq [a, b] \setminus A_1 = A_2$. If $c \in]x, y[$ then if $\varepsilon = \min(y - c, b - c) > 0$ then for $c + \frac{\varepsilon}{2}$ we have $c < c + \frac{\varepsilon}{2} < c + \varepsilon \leqslant y$, $b \Rightarrow c + \frac{\varepsilon}{2} \in]c, y[\bigcap]c, b[\subseteq]x, y[\bigcap A_2$ proving that $]x, y[\bigcap A_2 \neq \emptyset$ and as A_2 is closed we must have that $c \in A_2$. So we have finally that $c \in A_1 \bigcap A_2 = \emptyset$ a contradiction. So our assumption that $[a, b]$ is not connected is wrong and thus $[a, b]$ is connected what we set out to prove. □

Theorem 12.441. *Let $\langle \mathbb{R}, || \rangle$ be the a normed space then the following are equivalent for $I \subseteq \mathbb{R}$*

1. *I is connected*
2. *I is a generalized interval*

Proof.

1. **(1 ⇒ 2)** We proceed by contradiction so assume that I is connected but not a generalized interval. Using the properties of a generalized interval (see 2.180) we have that $\exists x, y \in I$ with $x < y$ and $\exists z \in \mathbb{R} \backslash I$ such that $x < z < y$ then $U_1 =]-\infty, z[\bigcap I$ contains x and is open in I, also $U_2 =]z, \infty[\bigcap I$ contains y and is open in I. Then as $z \in I$ we have $I \subseteq]-\infty, \infty[\backslash \{z\} =]-\infty, z[\bigcup]z, \infty[\Rightarrow I = I \bigcap (]-\infty, z[\bigcup]z, \infty[) = U_1 \bigcup U_2$, as also $U_1 \bigcap U_2 \subseteq]-\infty, z[\bigcap]z, \infty[= \emptyset \Rightarrow U_1 \bigcap U_2 = \emptyset$ we have that I is connected, contradicting our assumption that I is not connected. I must thus be connected.
2. **(2 ⇐ 1)** We prove this also by contradiction. So assume that I is a generalized interval and that I is not connected. Then we have open sets $U_1, U_2 \subseteq \mathbb{R}$ such that $I = (U_1 \bigcap I) \bigcup (U_2 \bigcap I)$, $\emptyset = (U_1 \bigcap I) \bigcap (U_2 \bigcap I) = U_1 \bigcap U_2 \bigcap I$ and $U_1 \bigcap I \neq \emptyset \neq U_2 \bigcap I$. So $\exists x \in U_1 \bigcap I$ and $\exists y \in U_2 \bigcap I$ and as $\emptyset = (U_1 \bigcap I) \bigcap (U_2 \bigcap I)$ we must have $x \neq y$, assume that $x < y$ (if $y < x$ interchange the role of U_1 and U_2) then using 2.180 we have $[x, y] \subseteq I \Rightarrow [x, y] = [x, y] \bigcap I = [x, y] \bigcap (U_1 \bigcap I) \bigcup (U_2 \bigcap I) = (U_1 \bigcap I \bigcap [x, y]) \bigcup (U_2 \bigcap I \bigcap [x, y]) = (U_1 \bigcap [x, y]) \bigcup (U_2 \bigcap [x, y]) \Rightarrow [x, y] = (U_1 \bigcap [x, y]) \bigcup (U_2 \bigcap [x, y])$, as also $x \in U_1 \bigcap [x, y]$, $y \in U_2 \bigcap [x, y]$ and $(U_1 \bigcap [x, y]) \bigcap (U_2 \bigcap [x, y]) \subseteq (U_1 \bigcap I) \bigcap (U_2 \bigcap I) = \emptyset$ we have that $[x, y]$ is covered by two non empty disjoint open sets and thus not connected, contradicting the fact that $[x, y]$ is connected by the previous theorem. □

Theorem 12.442. *Let $\langle X, \mathcal{T}_X \rangle$ be a connected topological space, $\langle Y, \mathcal{T}_Y \rangle$ a topological space and $f: X \rightarrow Y$ a continuous space then $f(X)$ is connected.*

Proof. We proceed by contradiction. Assume that $f(X)$ is not connected then we have open sets V_1, V_2 in Y such that $f(X) = (f(X) \bigcap V_1) \bigcup (f(X) \bigcap V_2) \subseteq V_1 \bigcup V_2$, $\exists y_1 \in f(X) \bigcap V_1$, $\exists y_2 \in f(X) \bigcap V_2$ and $(f(X) \bigcap V_1) \bigcap (f(X) \bigcap V_2) = \emptyset$. The $\exists x_1$, $x_2 \in X$ such that $f(x_1) = y_1 \in V_1$, $f(x_2) = y_2 \in V_2$ thus $x_1 \in f^{-1}(V_1) = U_1$, $x_2 \in f^{-1}(V_2) = U_2$ where U_1, U_2 are open because of continuity of f. Further if $x \in X \Rightarrow f(x) \in f(X) \subseteq V_1 \bigcup V_2 \Rightarrow x \in f^{-1}(V_1 \bigcup V_2) \subseteq f^{-1}(V_1) \bigcup f^{-1}(V_2) = U_1 \bigcup U_2$ proving that $X \subseteq U_1 \bigcup U_2$, as trivially $U_1 \bigcup U_2 \subseteq X$ we have $X = U_1 \bigcup U_2$. Finally assume that $y \in U_1 \bigcap U_2 \Rightarrow f(y) \in f(U_1) \bigcap f(U_2) \subseteq V_1 \bigcap V_2 = \emptyset$ a contradiction so that we have $U_1 \bigcap U_2 = \emptyset$. So we have proved that X is covered by non-empty disjoint open sets contradicting the fact that X is connected. □

Theorem 12.443. *Let $f: [a, b] \rightarrow \mathbb{R}$ a continuous function then $f([a, b]) = [c, d]$ (so f retains its minimum and maximum settings).*

Proof. As f is continuous and $[a, b]$ is compact we have by 12.243 that $f([a, b])$ is compact and thus by 12.242 it is bounded. As by the previous theorem $f([a, b])$ is connected we have by 12.441 that $f([a, b])$ is a generalized interval and as it is bounded there exists a $c, d \in \mathbb{R}$ such that $f([a, b]) = [c, d]$. □

Corollary 12.444. (Intermediate Value Theorem) *Let $f:[a,b]\rightarrow\mathbb{R}$ be a continuous function then if y is such that either $f(a)<y<f(b)$ or $f(b)<y<f(a)$ we have that there exists a $x\in[a,b]$ such that $f(x)=y$*

Proof. Using the previous theorem we have that $f([a,b])=[c,d]$ so that $f(a)$, $f(b)\in[c,d]$

$\boldsymbol{f(a)<y<f(b)}$**.** then $y\in[f(a),f(b)]$ and $f(a)<f(b)$ where $f(a),f(b)\in[c,d]$ so by the proeprties of the generalized interval (see 2.180 (2)) we have that $[f(a),f(b)]\subseteq[c,d]=f([a,b])$ hence $y\in f([a,b])$ and thus $\exists x\in[a,b]$ such that $y=f(x)$.

$\boldsymbol{f(b)<y<f(a)}$**.** then $y\in[f(b),f(a)]$ and $f(b)<f(a)$ where $f(a),f(b)\in[c,d]$ so by the proeprties of the generalized interval (see 2.180 (2)) we have that $[f(b),f(a)]\subseteq[c,d]=f([a,b])$ hence $y\in f([a,b])$ and thus $\exists x\in[a,b]$ such that $y=f(x)$. □

Chapter 13

Fundamental theorem of algebra and the Spectral theorems

13.1 Fundamental theorem of algebra

13.1.1 Polynomials

To prove the fundamental theorem of algebra which says that every non constant complex polynomial has a zero we need first a lot of definitions and lemmas.

Definition 13.1. *A function $p: \mathbb{C} \rightarrow \mathbb{C}$ is a complex polynomal if there exists a $n \in \mathbb{N}_0$ and a $\{a_i\}_{i \in \{0,\ldots,n\}} \subseteq \mathbb{C}$ such that $p(z) = \sum_{i=0}^{n} a_i \cdot z^i$. The set of complex ponynomials is noted by $\mathcal{P}$*

Next we prove that complex polynomials are uniquely defined by its coefficients.

Lemma 13.2. *Let $p: \mathbb{C} \rightarrow \mathbb{C}$ be a polynomial with coefficients $\{a_i\}_{i \in \{1,\ldots,n\}}, n \in \mathbb{N}_0$ such that $\forall x \in \mathbb{R} \subseteq \mathbb{C}$ we have $p(x) = 0$ then $\forall i \in \{1, \ldots, n\}$ we have $a_i = 0$*

Proof. Assume that $\exists i \in \{0, \ldots, n\}$ such that $a_i \neq 0$ then $\mathcal{N} = \{i \in \{0, \ldots, n\} | a_i \neq 0\} \neq \emptyset$ so that $m = \max(\mathcal{N})$ and $\forall x \in \mathbb{R}$ we have $0 = p(x) = \sum_{i=1}^{n} \alpha_i \cdot x^i = \sum_{i=1}^{m} a_i \cdot x^i$ and $a_m \neq 0$. We have then for m the following cases to consider

$m = 0$. then $0 = p(1) = a_0$ a contradiction

$0 < m$. then let $x = \frac{\sum_{i=0}^{m-1} |a_i|}{|a_m|} + 1$ (defined as $a_m \neq 0$) then $1 \leqslant x$ and thus by 9.31 we have $\forall j \in \{0, \ldots, m-1\}$ we have $x^j \leqslant x^{m-1}$, so that

$$\begin{aligned}
\left|\sum_{i=0}^{m-1} a_i \cdot x^i\right| &\leqslant_{12.70} \sum_{i=0}^{m-1} |a_i \cdot x^i| \\
&= \sum_{i=0}^{m-1} |a_i| \cdot |x|^i \\
&\leqslant \sum_{i=0}^{m-1} |a_i| \cdot |x|^{m-1} \\
&= |x|^{m-1} \cdot \sum_{i=0}^{m-1} |a_i| \\
&= |x|^{m-1} \cdot |a_m| \cdot (x-1) \\
&< |x|^m \cdot |a_m|
\end{aligned}$$

hence $\sum_{i=0}^{m-1} a_i \cdot x^i \neq -a_m \cdot x^m \Rightarrow \sum_{i=0}^{m} a_i \cdot x^i = \sum_{i=0}^{m-1} a_i \cdot x^i + a_m \cdot x^m \neq 0$ contradicting $\sum_{i=0}^{m} a_i \cdot z^i = 0$.

As in all cases we have a contradiction the assumption is wrong and thus $\forall i \in \{0, \ldots, n\}$ we have $a_i = 0$ □

Corollary 13.3. *Let $p: \mathbb{C} \rightarrow \mathbb{C}$ be a polynomial such that $\forall z \in \mathbb{C}$ we have that $p(z) = \sum_{i=0}^{n} a_i \cdot z^i = \sum_{i=0}^{n} b_i \cdot z^i$ then $\forall i \in \{0, \ldots, n\}$ we have $a_i = b_i$*

Proof. As $\forall z \in \mathbb{C}$ we have that $\sum_{i=0}^{n} a_i \cdot z^i = \sum_{i=0}^{n} b_i \cdot z^i \Rightarrow \sum_{i=1}^{n} (a_i - b_i) \cdot z^i = 0$ we have by the previous lemma that $\forall i \in \{0, \ldots, n\}$ we have $a_i - b_i \Rightarrow a_i = b_i$ □

Definition 13.4. *Let $p: \mathbb{C} \rightarrow \mathbb{C}$ be a polynomial defined by $p(x) = \sum_{i=1}^{n} a_i \cdot x^i$ then if we define $\mathcal{A} = \{i | a_i \neq 0\}$ we define* $\operatorname{ord}(p)$ *by* $\operatorname{ord}(p) = \begin{cases} 0 \text{ if } \mathcal{A} = \emptyset \\ \max(\mathcal{A}) \text{ if } \mathcal{A} \neq \emptyset \end{cases}$

Theorem 13.5. *Given the set of polynomials then we have that* $\operatorname{ord}: \mathcal{P} \rightarrow \mathbb{N}_0$ *defined by* $\operatorname{ord}(p)$ *is a function. Furthermore there exists a unique $\{c_i\}_{i \in \{0, \ldots, \operatorname{ord}(p)\}}$ such that $\forall z \in \mathbb{C}$ we have $p(z) = \sum_{i=0}^{\operatorname{ord}(p)} c_i \cdot z^i$*

Proof. We must proof that the value of the function is independent of the representation p, so let $p: \mathbb{C} \rightarrow \mathbb{C}$ be such that $p(z) = \sum_{i=0}^{n} a_i \cdot z^i = \sum_{i=0}^{m} b_i \cdot z^i$ (where we may assume witout loosing generality that $n \leqslant m$) then if we define $\{a_i'\}_{i \in \{1, \ldots, m\}}$ by $a_i' = \begin{cases} a_i \text{ if } i \in \{0, \ldots, n\} \\ 0 \text{ if } i \in \{n+1, \ldots, m\} \end{cases}$ then $\forall z \in \mathbb{C}$ we have $p(z) = \sum_{i=0}^{m} a_i' \cdot z^i = \sum_{i=0}^{n} b_i \cdot z^i$. Using 13.3 we conclude then that $\forall i \in \{1, \ldots, m\}$ we have $a_i' = b_i$ hence $\forall i \in \{1, \ldots, n\}$ we have $a_i = b_i$ and $\forall i \in \{n+1, \ldots, m\}$ we have $b_i = 0$. Hence $\{i \in \{0, \ldots, m\} | a_i \neq 0\} = \{i \in \{0, \ldots, n\} | b_i \neq 0\}$ so that $\min(\{i \in \{0, \ldots, m\} | a_i \neq 0\}) = \min(\{i \in \{0, \ldots, n\} | b_i \neq 0\})$ proving that $\operatorname{ord}(p)$ is independent of the representation of p. Fort he last part take $\{c_i\}_{i \in \{0, \ldots, \operatorname{ord}(p)\}}$ defined by $c_i = a_i$ for every $i \in \{0, \ldots, \operatorname{ord}(p)\}$. □

Definition 13.6. *A polynomial $p \in \mathcal{P}$ is **non constant** if* $\operatorname{ord}(p) \in \mathbb{N}$ *(in other words if there exists a $\{a_i\}_{i \in \{1, \ldots, n\}}$ with $a_n \neq 0$ such that $\forall z \in \mathbb{C}$ we have $p(z) = \sum_{i=0}^{n} a_i \cdot z^i$*

We show next that the product of a polynomial is again a polynomial, first we need a little lemma.

Lemma 13.7. *Let $n, m \in \mathbb{N}$ then $\{0, \ldots, n\} \times \{0, \ldots, m\} = \bigcup_{i \in \{0, \ldots, n+m\}} \{(r, s) \in \{0, \ldots, n\} \times \{0, \ldots, m\} | r + s = i\}$ and $\forall i, j \in \{0, \ldots, n+m\}$ with $i \neq j$ we have $\{(r, s) \in \{0, \ldots, n\} \times \{0, \ldots, m\} | r + s = i\} \bigcap \{(r, s) \in \{0, \ldots, n\} \times \{0, \ldots, m\} | r + s = j\} = \emptyset$*

Proof. Given $i \in \{0, \ldots, n+m\}$ define $I_i = \{(r, s) \in \{0, \ldots, n\} \times \{0, \ldots, m\} | r + s = i\} \subseteq \{0, \ldots n\} \times \{0, \ldots, m\}$. So trivially we have that $\bigcup_{i \in \{0, \ldots, n+m\}} I_i \subseteq \{0, \ldots, n\} \times \{0, \ldots, m\}$, further $(k, l) \in \{0, \ldots, n\} \times \{0, \ldots, m\}$ then $k + l \leqslant n + m$ hence $(k, l) \in I_{k+l} \subseteq \bigcup_{i \in \{0, \ldots, n+m\}} I_i$ proving that

$$\{0, \ldots, n\} \times \{0, \ldots, m\} = \bigcup_{i \in \{0, \ldots, n+m\}} \{(r, s) \in \{0, \ldots, n\} \times \{0, \ldots, m\} | r + s = i\}$$

Further if $i, j \in \{0, \ldots, n+m\}$ with $i = j$ we have for $(r, s) \in \{(r, s) \in \{0, \ldots, n\} \times \{0, \ldots, m\} | r+s = i\} \bigcap \{(r, s) \in \{0, \ldots, n\} \times \{0, \ldots, m\} | r+s = j\}$ we have $r+s = i \neq j = r+s$ a contradiction so we have that $\{(r, s) \in \{0, \ldots, n\} \times \{0, \ldots, m\} | r+s = i\} \bigcap \{(r, s) \in \{0, \ldots, n\} \times \{0, \ldots, m\} | r+s = j\} = \emptyset$.

□

Theorem 13.8. *Let $p, q \in \mathcal{P}$ be two polynomials defined by $p(z) = \sum_{i=0}^{n} a_i \cdot z^i$ and $q(z) = \sum_{i=0}^{m} b_i \cdot z^i$ then $p \cdot q: \mathbb{C} \to \mathbb{C}$ defined by $(p \cdot q)(z) = p(z) \cdot q(z)$ is also a polynomial defined by $(p \cdot q)(z) = \sum_{i=0}^{n+m} \left(\sum_{(k,l) \in \{(r,s) \in \{0,\ldots,n\} \times \{0,\ldots,m\} | r+s=i\}} a_k \cdot b_l\right) \cdot z^n$. Further if $\operatorname{ord}(p) = n \in \mathbb{N}$ and $\operatorname{ord}(q) = m \in \mathbb{N}$ then $\operatorname{ord}(p \cdot q) = n+m \in \mathbb{N}$*

Proof. Let $z \in \mathbb{C}$ then

$$\begin{aligned}
p(z) \cdot q(z) &= \left(\sum_{i=0}^{n} a_i \cdot z^i\right) \cdot \left(\sum_{j=0}^{m} b_i \cdot z^i\right) \\
&= \left(\sum_{i \in \{0,\ldots,n\}} a_i \cdot z^i\right) \cdot \left(\sum_{j \in \{0,\ldots,m\}} b_i \cdot z^i\right) \\
&\underset{10.61}{=} \sum_{(i,j) \in \{0,\ldots,n\} \times \{0,\ldots,m\}} a_i \cdot b_i \cdot z^i \cdot z^j \\
&= \sum_{(i,j) \in \{0,\ldots,n\} \times \{0,\ldots,m\}} a_i \cdot b_i \cdot z^{i+j} \\
&\underset{10.47 \text{ togehter with } 13.7}{=} \sum_{i \in \{0,\ldots,n+m\}} \left(\sum_{(k,l) \in \{(r,s) \in \{0,\ldots,n\} \times \{0,\ldots,m\} | r+s=i\}} a_k \cdot b_l \cdot z^{k+l}\right) \\
&= \sum_{i \in \{0,\ldots,n+m\}} \left(\sum_{(k,l) \in \{(r,s) \in \{0,\ldots,n\} \times \{0,\ldots,m\} | r+s=i\}} a_k \cdot b_l \cdot z^i\right) \\
&= \sum_{i \in \{0,\ldots,n+m\}} \left(\sum_{(k,l) \in \{(r,s) \in \{0,\ldots,n\} \times \{0,\ldots,m\} | r+s=i\}} a_k \cdot b_l\right) \cdot z^i \\
&= \sum_{i=0}^{n+m} \left(\sum_{(k,l) \in \{(r,s) \in \{0,\ldots,n\} \times \{0,\ldots,m\} | r+s=i\}} a_k \cdot b_l\right) \cdot z^i
\end{aligned}$$

Finally if $\operatorname{ord}(p), \operatorname{ord}(q) \in \mathbb{N}$ then $a_n, a_m \neq 0$ hence $a_n \cdot a_m \neq 0$, as $\{(r, s) \in \{0, \ldots, n\} \times \{0, \ldots, m\} | r+s = n+m\} = \{(n, m)\}$ we have that the coefficent of z^{n+m} is $a_n \cdot a_m \neq 0$ hence $\operatorname{ord}(p \cdot q) = n+m$. □

Definition 13.9. *Let $n \in \mathbb{N}$ and $\{p_i\}_{i \in \{1,\ldots,n\}} \subseteq \mathcal{P}$ a finite non empty family of polynomials then $\prod_{i=1}^{n} p_i: \mathbb{C} \rightarrow \mathbb{C}$ is defined by $(\prod_{i=1}^{n} p_i)(z) = \prod_{i=1}^{n} p_i(z)$*

Using mathematical induction we can then prove that $\prod_{i=1}^{n} p_i$ is also a polynomial

Lemma 13.10. *Let $n \in \mathbb{N}$ and $\{p_i\}_{i \in \{1,\ldots,n\}} \subseteq \mathcal{P}$ a finite non empty family of polynomials then $\prod_{i=1}^{n} p_i: \mathbb{C} \rightarrow \mathbb{C}$ is a polynomial with $\operatorname{ord}(\prod_{i=1}^{n} p_i) = \sum_{i=1}^{n} \operatorname{ord}(p_i)$*

Proof. Let $\mathcal{S} = \{n \in \mathbb{N} | \text{if } \{p_i\}_{i \in \{1,\ldots,n\}} \subseteq \mathcal{P} \text{ then } \prod_{i=1}^{n} p_i: \mathbb{C} \rightarrow \mathbb{C} \text{ is a polynomial with } \operatorname{ord}(\prod_{i=1}^{n} p_i) = \sum_{i=1}^{n} \operatorname{ord}(p_i)\}$ then we have

$1 \in \mathcal{S}$. then $\forall z \in \mathbb{C}$ we have $(\prod_{i=1}^{1} p_i)(z) = p_1(z)$ proving that $\prod_{i=1}^{1} p_i = p_1$ a polynomial and $\operatorname{ord}(\prod_{i=1}^{1} p_i) = \operatorname{ord}(p_1) = \sum_{i=1}^{1} \operatorname{ord}(p_i)$ hence $1 \in \mathcal{S}$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$. then $\forall z \in \mathbb{C}$ we have that $(\prod_{i=1}^{n+1} p_i)(x) = \prod_{i=1}^{n+1} p_i(z) = p_{n+1}(z) \cdot (\prod_{i=1}^{n} p_i)(z)$ proving that $\prod_{i=1}^{n+1} p_i = p_{n+1} \cdot \prod_{i=1}^{n} p_i$ which as $n \in \mathcal{S}$ is a product of polynomials and thus by 13.8 is a polynomial, further

$$\begin{aligned}
\operatorname{ord}\left(\prod_{i=1}^{n+1} p_i\right) &= \operatorname{ord}\left(p_{n+1} \cdot \prod_{i=1}^{n} p_i\right) \\
&\underset{13.8}{=} \operatorname{ord}(p_{n+1}) + \operatorname{ord}\left(\prod_{i=1}^{n} p_i\right) \\
&\underset{n \in \mathcal{S}}{=} \operatorname{ord}(p_{n+1}) + \sum_{i=1}^{n} \operatorname{ord}(p_i) \\
&= \sum_{i=1}^{n+1} \operatorname{ord}(p_i)
\end{aligned}$$

hence $n+1 \in \mathcal{S}$ □

13.1.2 Divergent sequences

We need also the concept of divergent limits to prove the fundamental theorem of algebra

Definition 13.11. *Let $f: \mathbb{R} \rightarrow \mathbb{R}$ be a function then*

1. $\lim_{x \rightarrow \infty} f(x) = \infty$ *if $\forall M \in \mathbb{R}$ there exists a $N \in \mathbb{R}$ such that $\forall x \geqslant N$ we have $f(x) \geqslant M$*
2. $\lim_{x \rightarrow \infty} f(x) = -\infty$ *if $\forall M \in \mathbb{R}$ there exists a $N \in \mathbb{R}$ such that $\forall x \geqslant N$ we have $f(x) \leqslant M$*

Note 13.12. Let $f: \mathbb{R} \rightarrow \mathbb{R}$ be a function and consider then $-f: \mathbb{R} \rightarrow \mathbb{R}$ defined by $(-f)(x) = -f(x)$

1. If $\lim_{x \rightarrow \infty} f(x) = \infty$ then $\lim_{x \rightarrow \infty} (-f)(x) = -\infty$

2. If $\lim_{x \to \infty} f(x) = -\infty$ then $\lim_{x \to \infty} (-f)(x) = \infty$

Proof.

1. Take $\mathrm{M} \in \mathbb{R}$ then there exists a N such that $\forall x \geqslant N$ we have $-M \leqslant f(x) \Rightarrow (-f)(x) \leqslant M$ proving that $\lim_{x \to \infty} (-f)(x) = -\infty$
2. Take $\mathrm{M} \in \mathbb{R}$ then there exists a N such that $\forall x \geqslant N$ we have $f(x) \leqslant -M \Rightarrow M \leqslant (-f)(x)$ proving that $\lim_{x \to \infty} (-f)(x) = \infty$ □

Lemma 13.13. *Let* $n \in \mathbb{N}$, $\{a_i\}_{i \in \{0, \ldots, n\}} \subseteq \mathbb{R}$ *then*

1. *If* $a_n > 0$ *we have* $\lim_{x \to \infty} (\sum_{i=0}^n a_i \cdot x^i) = \infty$
2. *If* $a_n < \infty$ *we have* $\lim_{x \to \infty} (\sum_{i=0}^n a_i \cdot x^i) = -\infty$

Proof.

1. we prove this by induction so let $\mathcal{S} = \left\{ n \in \mathbb{N} | \forall \{a_i\}_{i \in \{0, \ldots, n\}} \subseteq \mathbb{R} \vdash a_n > 0 \text{ we have } \lim_{x \to \infty} (\sum_{i=0}^n a_i \cdot x^i) = \infty \right\}$ then we have

   $1 \in \mathcal{S}$. Let $M \in \mathbb{R}$ and take $N = \frac{M - a_0}{a_1}$ then if $x \geqslant N$ we have $M - a_0 = a_1 \cdot N \leqslant a_1 \cdot x = a_1 \cdot x_1$ hence $M \leqslant a_1 \cdot x^1 + a_0 = \sum_{i=0}^1 a_i \cdot x^i$ proving that $1 \in \mathcal{S}$

   $n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$. Let $\{a_i\}_{i \in \{0, \ldots, n+1\}} \subseteq \mathbb{R} \vdash a_{n+1} > 0$ then $\sum_{i=0}^{n+1} a_i \cdot x^i = (\sum_{i=1}^{n+1} a_i \cdot x^i) + a_0 = x \cdot (\sum_{i=1}^{n+1} a_i \cdot x^{i-1}) + a_0 = x \cdot (\sum_{i=0}^n a_{i+1} \cdot x^i) + a_0$, hence if we define $\{b_i\}_{i \in \{0, \ldots, n\}}$ by $b_i = a_{i+1}$, then we have

   $$\sum_{i=1}^{n+1} a_i \cdot x^i = x \cdot \left(\sum_{i=0}^n b^i \cdot x^i \right) + a_0 \wedge b_n = a_{n+1} > 0$$

   Take $M \in \mathbb{R}$ then as $n \in \mathcal{S}$ there exists a $N' \in \mathbb{R}$ such that if $N \leqslant x$ we have that $\max(1, M - a_0) \leqslant \sum_{i=0}^n b^i \cdot x^i$. Take then $N = \max(N', 1)$ then we have $1 \leqslant x$ and as $0 < \max(1, M - a_0)$ we have $\max(1, M - a_0) \leqslant x \cdot \max(1, M - a_0)$. As $0 < 1 \leqslant x$ then we have $x \cdot \max(1, M - a_0) \leqslant x \cdot (\sum_{i=0}^n b^i \cdot x^i)$. So we conclude that $M - a_0 = \max(1, M - a_0) \leqslant x \cdot (\sum_{i=0}^n b^i \cdot x^i)$ and thus $M \leqslant \cdot (\sum_{i=0}^n b^i \cdot x^i) + a_0 = \sum_{i=1}^{n+1} a_i \cdot x^i$. This proves that $n + 1 \in \mathcal{S}$
2. If $a_n < 0$ then $0 < (-a_n)$ and we have by (1) that $\lim_{x \to \infty} (\sum_{i=0}^n (-a_i) \cdot x^i) = \infty$ and thus by 13.12 we have that $\lim_{x \to \infty} (\sum_{i=0}^n a_i \cdot x^i) = \lim_{x \to \infty} (-(\sum_{i=0}^n a_i \cdot x^i)) = -\infty$ □

13.1.3 Properties of $\mathbb{C}$ needed for the fundamental theorem

We define now a norm on $\mathbb{C}$ that we will use instead of the classical norm on $\mathbb{C}$ that includes a square root.

Lemma 13.14. *$\|\|:\mathbb{C}\rightarrow\mathbb{R}$ by $\|z\|=|\mathrm{Re}(z)|+|\mathrm{Img}(z)|$ is a norm*

Proof. We have

1. If $z\in\mathbb{C}$ then trivially $0\leqslant\|z\|$
2. If $\alpha\in\mathbb{R}$ then for $z=x+i\cdot y$ we have that $\|\alpha\cdot z\|=\|\alpha\cdot x+i\cdot\alpha\cdot y\|=|\alpha\cdot x|+|\alpha\cdot y|=|\alpha|\cdot(|x|+|y|)=\alpha\cdot\|z\|$
3. If $z_1=x_1+i\cdot y_1,\ z_2=x_2+i\cdot y_2\in\mathbb{C}$ then $\|z_1+z_2\|=\|(x_1+i\cdot y_1)+(x_2+i\cdot y_2)\|=\|(x_1+x_2)+i\cdot(y_1+y_2)\|=|x_1+x_2|+|y_1+y_2|\leqslant|x_1|+|x_2|+|y_1|+|y_2|=\|z_1\|+\|z_2\|$
4. Take $z=z+i\cdot y\in\mathbb{C}$ such that $\|z\|=0$ then $|x|+|y|=0\underset{0\leqslant|x|,|u|}{\Rightarrow}|x|=0=|y|\Rightarrow x=0\wedge y=0\Rightarrow z=0$ ☐

The above norm has the following properties

Lemma 13.15. *We have the following properties for the above norm*

1. *If $z\in\mathbb{R}$ then $\|z\|=|z|$*
2. *$\forall z\in\mathbb{C}$ we have that $\|\bar{z}\|=\|z\|$*
3. *$\forall z,w\in\mathbb{C}$ we have that $\frac{\|z\|\cdot\|w\|}{2}\leqslant\|z\cdot w\|\leqslant\|z\|\cdot\|w\|$*
4. *$\forall z,w\in\mathbb{C}$ we have $\|\mathrm{Re}(z\cdot w)\|\leqslant\|z\|\cdot\|w\|$*
5. *$\forall n\in\mathbb{N},\ z\in\mathbb{C}$ then $\|z^n\|\leqslant\|z\|^n$*
6. *$\forall n\in\mathbb{N},\ z\in\mathbb{C}$ then $\|\mathrm{Re}(z^n)\|\leqslant\|z\|^n$*
7. *$\forall n\in\mathbb{N},z\in\mathbb{C}$ then $\frac{\|z\|^n}{2^n}\leqslant\|z^n\|$*

Proof.

1. If $z\in\mathbb{R}$ then $\mathrm{Img}(z)=0$ and $\mathrm{Re}(z)=z$ thus $\|z\|=|\mathrm{Re}(z)|+|\mathrm{Img}(z)|=|z|+0=|z|$
2. If $z=x+i\cdot y$ then we have $\|z\|=|x|+|y|=|x|+|-y|=\|x-i\cdot y\|=\|\bar{z}\|$
3. First we have given $a,b,c,d\in\mathbb{R}$ that $0\leqslant(|a|-|b|)^2=a^2+b^2-2\cdot|a|\cdot|b|$ proving that $2\cdot|a|\cdot|b|\leqslant a^2+b^2$ hence $(|a|+|b|)^2=a^2+b^2+2\cdot|a|\cdot|b|\leqslant 2\cdot(a^2+b^2)$. Similar we have that $0\leqslant(|c|-|d|)^2=c^2+d^2-2\cdot|c|\cdot|d|$ proving that $2\cdot|c|\cdot|d|\leqslant c^2+d^2$ hence $(|c|+|d|)^2=c^2+d^2+2\cdot|c|\cdot|d|\leqslant 2\cdot(c^2+d^2)$. So $(|a|+|b|)^2\cdot(|c|+|d|)^2\leqslant 2\cdot(a^2+b^2)\cdot(|c|+|d|)^2$ and $2\cdot(|c|+|d|)^2\cdot(a^2+b^2)\leqslant 4\cdot(c^2+d^2)$ proving that
$$(|a|+|b|)^2\cdot(|c|+|d|)^2\leqslant 4\cdot(a^2+b^2)\cdot(c^2+d^2) \tag{13.1}$$
Now
$$\begin{aligned}(a\cdot c-b\cdot d)^2+(a\cdot d+b\cdot c)^2 &= a^2\cdot c^2+b^2\cdot d^2-2\cdot a\cdot c\cdot b\cdot d+a^2\cdot d^2+b^2\cdot c^2+\\ &\quad 2\cdot a\cdot d\cdot b\cdot c\\ &= a^2\cdot c^2+a^2\cdot d^2+b^2\cdot d^2+b^2\cdot c^2\\ &= a^2\cdot(c^2+d^2)+b^2\cdot(c^2+d^2)\\ &= (a^2+b^2)\cdot(c^2+d^2)\end{aligned}$$

Using 13.1 we have then

$$\begin{array}{rcl}(|a|+|b|)^2\cdot(|c|+|d|)^2 & \leqslant & 4\cdot((a\cdot c-b\cdot d)^2+(a\cdot d+b\cdot c)^2)\\ & \leqslant & 4\cdot((a\cdot c-b\cdot d)^2+(a\cdot d+b\cdot c)^2+2\cdot(a\cdot c-b\cdot d)\cdot\\ & & (a\cdot d+b\cdot c))\\ & = & 4\cdot((a\cdot c-b\cdot d)+(a\cdot d+b\cdot c))^2\\ & = & 4\cdot|(a\cdot c-b\cdot d)+(a\cdot d+b\cdot c)|^2\end{array}$$

proving taking the square root (see 9.71) $(|a|+|b|)\cdot(|c|+|d|)\leqslant 2\cdot|(a\cdot c-b\cdot d)+(a\cdot d+b\cdot c)|\leqslant|a\cdot c-b\cdot d|+|a\cdot d+b\cdot c|$

$$(|a|+|b|)\cdot(|c|+|d|)\leqslant 2\cdot|a\cdot c-b\cdot d|+|a\cdot d+b\cdot c| \tag{13.2}$$

So if $z=a+i\cdot b$ and $w=c+i\cdot d$ then $z\cdot w=(a\cdot c-b\cdot d)+i\cdot(a\cdot d+b\cdot c)$ so that $\|z\cdot w\|=|a\cdot c-b\cdot d|+|a\cdot d+b\cdot c|\geqslant_{13.2}\frac{1}{2}\cdot(|a|+|b|)\cdot(|c|+|d|)=\frac{1}{2}\cdot\|z\|\cdot\|w\|$ proving that

$$\frac{1}{2}\cdot\|z\|\cdot\|w\|\leqslant\|z\cdot w\| \tag{13.3}$$

Further we have

$$\begin{array}{rcl}|a\cdot c-b\cdot d|+|a\cdot d+b\cdot c| & \leqslant & |a|\cdot|c|+|b|\cdot|d|+|a|\cdot|d|+|b|\cdot|c|\\ & = & |a|\cdot(|c|+|d|)+|b|\cdot(|c|+|d|)\\ & = & (|a|+|b|)\cdot(|c|+|d|)\end{array}$$

proving

$$|a\cdot c-b\cdot d|+|a\cdot d+b\cdot c|\leqslant(|a|+|b|)\cdot(|c|+|d|) \tag{13.4}$$

hence

$$\begin{array}{rcl}\|z\cdot w\| & = & |a\cdot c-b\cdot d|+|a\cdot d+b\cdot c|\\ & \leqslant_{13.4} & (|a|+|b|)\cdot(|c|+|d|)\\ & = & \|z\|\cdot\|w\|\end{array}$$

proving

$$\|z\cdot w\|\leqslant\|z\|\cdot\|w\| \tag{13.5}$$

(3) is thus proved by 13.3 and 13.5.

4. If $z=a+i\cdot b$ and $w=c+i\cdot d$ then

$$\begin{array}{rcl}\|\mathrm{Re}(z\cdot w)\| & \underset{(1)}{=} & |\mathrm{Re}((a\cdot c-b\cdot d)+(a\cdot d+b\cdot c))|\\ & = & |a\cdot c-b\cdot d|\\ & \leqslant & |a\cdot c-b\cdot d|+|a\cdot d+b\cdot c|\\ & = & \|z\cdot w\|\\ & \leqslant_{(3)} & \|z\|\cdot\|w\|\end{array}$$

5. This trivially proved by induction so let $\mathcal{S}=\{n\in\mathbb{N}|\|z^n\|\leqslant\|z\|^n\}$ then we have

   $1\in\mathcal{S}$. then $\|z^1\|=\|z\|=\|z\|^1$ proving that $1\in\mathcal{S}$

$n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$. then $\|z^{n+1}\| = \|z \cdot z^n\| \leqslant_{(3)} \|z\| \cdot \|z\|^n = \|z\|^{n+1}$ proving that $n+1\in\mathcal{S}$

6. Take $z = x + i\cdot y$ then we must consider two cases for n

   $n=1$. then $\|\mathrm{Re}(z^1)\| = \|x\| \underset{(1)}{=} |x| \leqslant |x| + |y| = \|z\|$

   $n>1$. then $1\leqslant n-1$ and $\|\mathrm{Re}(z^n)\| = \|\mathrm{Re}(z\cdot z^{n-1})\| \leqslant \|z\| \cdot \|z^{n-1}\| \leqslant_{(3)} \|z\|\cdot\|z\|^{n-1} = \|z\|^n$

7. This is proved by induction so let $\mathcal{S} = \left\{n\in\mathbb{N}_0 \,|\, \frac{\|z\|^n}{2^n} \leqslant \|z^n\|\right\}$ then we have

   $1\in\mathcal{S}$. then as $\frac{1}{2}<1\Rightarrow\frac{\|z\|}{2^1} = \frac{1}{2}\cdot\|z\| \leqslant 1\cdot\|z\| = \|z\|^1$ we have that $1\in\mathcal{S}$

   $n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$. we have $\frac{\|z\|\cdot\|z^n\|}{2} \leqslant_{(3)} \|z\cdot z^n\| = \|z^{n+1}\|$ as $n\in\mathcal{S}$ we have $\frac{\|z\|^n}{2^n} \leqslant \|z^n\|$ and thus $\frac{\|z\|^{n+1}}{2^{n+1}} = \frac{\|z\|}{2}\cdot\frac{\|z\|^n}{2^n} \leqslant \frac{\|z\|}{2}\cdot\|z^n\| \leqslant \|z^{n+1}\|$ proving that $n+1\in\mathcal{S}$ □

In the proof of the fundamental theorem of algebra we use the extrem value theorem so we need continuity using the above norm. This is proved in the following theorem.

Theorem 13.16. *The following functions are continuous in the normed space* $\langle\mathbb{C},\|\|\rangle$

1. *Given $\alpha\in\mathbb{C}$ then function $(\alpha\cdot):\mathbb{C}\rightarrow\mathbb{C}$ defined by $(\alpha\cdot)(x)=\alpha\cdot x$ is continuous*
2. Given $n\in\mathbb{N}_0$ *we have that $(1_{\mathbb{C}})^n:\mathbb{C}\rightarrow\mathbb{C}$ is continuous*
3. *Given $n\in\mathbb{N}$, $\{\alpha_i\}_{i\in\mathbb{N}_0\{0,\ldots,n\}}$ the function $p:\mathbb{C}\rightarrow\mathbb{C}$ defined by $p(x)=\sum_{i=0}^n \alpha_i\cdot x^i$ is continuous*
4. *The function $\langle\rangle:\mathbb{C}\rightarrow\mathbb{R}$ defined by $\langle z\rangle = z\cdot\bar{z} = |z|^2$ is continuous*
5. *Given $n\in\mathbb{N}$, $\{a_i\}_{i\in\mathbb{N}}$ the function $p\cdot\bar{p}:\mathbb{C}\rightarrow\mathbb{R}$ defined by $(p\cdot\bar{p})(x) = \left(\sum_{i=0}^n\alpha_i\cdot x^i\right)\cdot\overline{\left(\sum_{i=0}^n\alpha_i\cdot x^i\right)}$ is continuous*
6. *The function $^-:\mathbb{C}\rightarrow\mathbb{C}$ defined by $^-(z)=\bar{z}$ is continuous*
7. *Given $z\in\mathbb{C}$ then the function $(\cdot z):\mathbb{R}\rightarrow\mathbb{C}$ defined by $(.z)(\alpha)=\alpha\cdot z$ is continuous*
8. *If $p:\mathbb{C}\rightarrow\mathbb{C}$ is a polynomial and $z\in\mathbb{C}$ then $p_{\cdot z}:\mathbb{R}\rightarrow\mathbb{C}$ defined by $p_{\cdot z}(\alpha) = p(\alpha\cdot z)$ is contnuous.*
9. *The function* Re: $\mathbb{C}\rightarrow\mathbb{R}$ *defined by $x\rightarrow\mathrm{Re}(x)$ is continuous*

Proof.

1. *As $(\alpha\cdot)=\alpha\cdot 1_{\mathbb{C}}$ this follows from the fact that $1_{\mathbb{C}}$ is continuous (see 12.137) and 12.170*
2. *As $1_{\mathbb{C}}$ is a continuous function (see 12.137) this follows from 12.173.*

3. *If* $\{a_i\}_{i\in\{0,\ldots,n\}}$ *we have that* $\forall i\in\{1,\ldots,n\}$ *that* $(\alpha_i\cdot)\circ(1_{\mathbb{C}})^n$ *is continuous as the composition of two continuous functions (see (1) and (2). Hence usng 12.171 we have that* $\sum_{i=1}^n(\alpha\cdot)\circ(1_{\mathbb{C}})^i$ *is continuous. As* $\forall x\in\mathbb{C}$ *we have* $(\sum_{i=1}^n(\alpha\cdot)\circ(1_{\mathbb{C}})^i)(x)=\sum_{i=1}^n((\alpha\cdot,)\circ(1_{\mathbb{C}})^i)=\sum_{i=1}^n\alpha_i\cdot x^i=p(x)$ *we have that* $p=\sum_{i=1}^n(\alpha\cdot)\circ(1_{\mathbb{C}})^i$ *is continuous.*

4. *First let* $z,\delta\in\mathbb{C}$ *then*

$$\begin{array}{rcl}|\langle z+\delta\rangle-\langle z\rangle| & = & \|\langle z+\delta\rangle-\langle z\rangle\|\\ & = & \|(z+\delta)\cdot\overline{(z+\delta)}-z\cdot\bar{z}\|\\ & = & \|(z+\delta)\cdot(\bar{z}+\bar{\delta})-z\cdot\bar{z}\|\\ & = & \|z\cdot\bar{z}+z\cdot\bar{\delta}+\delta\cdot\bar{z}+\delta\cdot\bar{\delta}-z\cdot\bar{z}\|\\ & = & \|z\cdot\bar{\delta}+\delta\cdot\bar{z}+\delta\cdot\bar{\delta}\|\\ & \leqslant & \|z\cdot\bar{\delta}\|+\|\delta\cdot\bar{z}\|+\|\delta\cdot\bar{\delta}\|\\ & \leqslant_{13.15\ (3)} & \|z\|\cdot\|\bar{\delta}\|+\|\delta\|\cdot\|\bar{z}\|+\|\delta\|\cdot\|\bar{\delta}\|\\ & \leqslant_{13.15\ (2)} & \|z\|\cdot\|\delta\|+\|\delta\|\cdot\|z\|+\|\delta\|\cdot\|\delta\|\\ & = & \|\delta\|\cdot(2\cdot\|z\|+\|\delta\|)\end{array}$$

So given $\varepsilon>0$ *we have if we choose* δ *such that* $\|\delta\|<\min\left(1,\frac{\varepsilon}{2\cdot\|z\|+1}\right)$ *we have that*

$$\begin{array}{rcl}\|\langle z+\delta\rangle-\langle z\rangle\| & < & \|\delta\|\cdot(2\cdot\|z\|+1)\\ & < & \dfrac{\varepsilon}{2\cdot\|z\|+1}\cdot(2\cdot\|z\|+1)\\ & = & \delta\end{array}$$

proving continuity of $\langle\rangle$.

5. *This follows as* $p\cdot\bar{p}=\langle\rangle\circ p$ *a composition of continuous functions by (3) and (4)*

6. *Let* $z\in\mathbb{C}$ *and take* $\varepsilon>0$ *then if* $z'\in\mathbb{C}$ *is such that* $\|z-z'\|<\varepsilon$ *we have that* $\|^-(z)-{}^-(z')\|=\|\bar{z}-\bar{z}'\|=\|\overline{z-z'}\|\underset{13.15(2)}{=}\|z-z'\|<\varepsilon$

7. *Let* $z\in\mathbb{C}$, $\alpha\in\mathbb{R}$, *given* $\varepsilon>0$ *define* $\delta=\frac{\varepsilon}{\|z\|+1}>0$ *then we have* $\forall\alpha'\in\mathbb{R}$ *with* $|\alpha-\alpha'|<\delta$ *that* $\|(\cdot z)(\alpha)-(\cdot z)(\alpha')\|=\|\alpha\cdot z-\alpha'\cdot z\|=\|(\alpha-\alpha')\cdot z\|=|\alpha-\alpha'|\cdot\|z\|<\frac{\varepsilon}{\|z\|+1}\cdot\|z\|<\varepsilon$ *proving that* $(\cdot z)$ *is continuous (see 12.152).*

8. *As* $p_{\cdot z}=p\circ(\cdot z)$ *is the composition of two continuous functions (see (3) and (7) we have that* $p_{\cdot z}$ *is continuous.*

9. *Let* $z\in\mathbb{C}$ *then given* ε *we have if* $\delta=\varepsilon$ *that* $\forall z'\in\mathbb{C}$ *with* $\|z-z'\|<\delta$ *that* $|\mathrm{Re}(z)-\mathrm{Re}(z')|=|\mathrm{Re}(z-z')|\leqslant|\mathrm{Re}(z-z')|+|\mathrm{Img}(z-z')|=\|z-z'\|<\delta=\varepsilon$ *proving continuity.* □

We need in the proof also to calculate the value of $(1+i)^{4\cdot n+2}$ and $(1-i)^{4\cdot n+2}$

Lemma 13.17. *Let* $n\in\mathbb{N}_0$ *then* $(1+i)^{4\cdot n+2}=2\cdot(-4)^n\cdot i$ *and* $(1-i)^{4\cdot n+2}=-2\cdot(-4)^n\cdot i$

Proof. We use induction to prove this

1. Take $\mathcal{S}=\{n\in\mathbb{N}_0 \,|\, (1+i)^{4\cdot n+2}=2\cdot(-4)^n\cdot i\}$ then we have

   $\boldsymbol{0\in\mathcal{S}}$. $(1+i)^{4\cdot 0+2}=(1+i)^2=1+2\cdot i+i^2=2\cdot i=2\cdot(-4)^0\cdot i$ proving that $0\in\mathcal{S}$

   $\boldsymbol{n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}}$.

$$\begin{array}{rcl}(1+i)^{4\cdot(n+1)+2} & = & (1+i)^{4\cdot n+2+4}\\ & = & (1+i)^{4\cdot n+2}(1+i)^4\\ & \underset{n\in\mathcal{S}}{=} & 2\cdot(-4)^n\cdot i\cdot(1+i)^2(1+i)^2\\ & = & 2\cdot(-4)^n\cdot i\cdot(2\cdot i)(2\cdot i)\\ & = & 2\cdot(-4)^n\cdot i\cdot(-4)\\ & = & 2\cdot(-4)^{n+1}\cdot i\end{array}$$

   proving that $n+1\in\boldsymbol{S}$

2. Take $\mathcal{S}=\{n\in\mathbb{N}_0 \,|\, (1-i)^{4\cdot n+2}=-2\cdot(-4)^n\cdot i\}$ then we have

   $\boldsymbol{0\in\mathcal{S}}$. $(1-i)^{4\cdot 0+2}=(1-i)^2=1-2\cdot i+(-i)^2=-2\cdot i=-2\cdot(-4)^0\cdot i$ proving that $0\in\mathcal{S}$

   $\boldsymbol{n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}}$.

$$\begin{array}{rcl}(1-i)^{4\cdot(n+1)+2} & = & (1-i)^{4\cdot n+2+4}\\ & = & (1-i)^{4\cdot n+2}(1-i)^4\\ & \underset{n\in\mathcal{S}}{=} & -2\cdot(-4)^n\cdot i\cdot(1-i)^2(1-i)^2\\ & = & -2\cdot(-4)^n\cdot i\cdot(-2\cdot i)(-2\cdot i)\\ & = & -2\cdot(-4)^n\cdot i\cdot(-4)\\ & = & -2\cdot(-4)^{n+1}\cdot i\end{array}$$

   proving that $n+1\in\boldsymbol{S}$ ☐

The binomial expansion of the power of a sum is used in the theorem so we have to define the factorial and the binomial constants and prove the binomial formula.

Definition 13.18. *Given $n\in\mathbb{N}_0$ we have that $n!$ is defined recursively by* $n!=\begin{cases}1 \text{ if } n=0\\ n\cdot(n-1)! \text{ if } n>1\end{cases}$

Definition 13.19. *Given $n\in\mathbb{N}_0$ and $k\in\{0,\ldots,n\}$ then* $\binom{n}{k}=\frac{n!}{k!\cdot(n-k)!}$

Note 13.20. If $n\in\mathbb{N}_0$ then $\binom{n}{0}=1$ and $\binom{n}{n}=1$

Proof. $\binom{n}{0}=\frac{n!}{0!\cdot(n-0)!}=\frac{n!}{1\cdot n!}=1$ and $\binom{n}{n}=\frac{n!}{n!\cdot(n-n)!}=\frac{1}{0!}=1$ ☐

Theorem 13.21. *If $n\in\mathbb{N}$ and $0<k\leqslant n$ then* $\binom{n}{k}+\binom{n}{k-1}=\binom{n+1}{k}$

Proof.

$$\begin{aligned}
\binom{n}{k}+\binom{n}{k-1} &= \frac{n!}{k!\cdot(n-k)!}+\frac{n!}{(k-1)!\cdot(n-(k-1))!}\\
&= \frac{n!\cdot((n+1)-k)}{k!\cdot(n-k)!\cdot((n+1)-k)}+\frac{n!\cdot k}{k\cdot(k-1)!\cdot(n+1-k)}\\
&= \frac{n!\cdot(n+1-k)}{k!\cdot((n+1)-k)!}+\frac{n!\cdot k}{k!\cdot((n+1)-k)!}\\
&= \frac{n!\cdot(n+1)}{k!\cdot((n+1)-k)!}\\
&= \frac{(n+1)!}{k!\cdot((n+1)-k)!}\\
&= \binom{n+1}{k}
\end{aligned}$$

□

Theorem 13.22. **(Binomial formula)** *Let F be a field, $n\in\mathbb{N}$, $a,b\in F$ then $(a+b)^n=\sum_{k=0}^{n}\left(\binom{n}{k}\cdot a^k\cdot b^{n-k}\right)$*

Proof. We prove this theorem by mathematical induction so let $\mathcal{S}=\left\{n\in\mathbb{N}|(a+b)^n=\sum_{k=0}^{n}\left(\binom{n}{k}\cdot a^k\cdot b^{(n-k)}\right)\right\}$ then we have

$1\in\mathcal{S}$. $\sum_{k=0}^{1}\left(\binom{1}{k}\cdot a^k\cdot b^{(n-k)}\right)=\binom{1}{0}\cdot a^0\cdot b^{(1-0)}+\binom{1}{1}\cdot a^1\cdot b^{(1-1)}\underset{13.20}{=}b+a=(a+b)^1$ proving that $1\in\mathcal{S}$.

$n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$.

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)^n\cdot(a+b)\\
&\underset{n\in\mathcal{S}}{=} \left(\sum_{k=0}^{n}\left(\binom{n}{k}\cdot a^k\cdot b^{(n-k)}\right)\right)\cdot(a+b)\\
&= \sum_{k=0}^{n}\left(\binom{n}{k}\cdot a^{k+1}\cdot b^{(n-k)}\right)+\sum_{k=0}^{n}\left(\binom{n}{k}\cdot a^k\cdot b^{(n-k+1)}\right)\\
&= \sum_{k=0}^{n}\left(\binom{n}{k}\cdot a^{k+1}\cdot b^{((n+1)-(k+1))}\right)+\sum_{k=0}^{n}\left(\binom{n}{k}\cdot a^k\cdot b^{((n+1)-k)}\right)\\
&= \sum_{k=1}^{n+1}\left(\binom{n}{k-1}\cdot a^k\cdot b^{((n+1)-k)}\right)+\sum_{k=0}^{n}\left(\binom{n}{k}\cdot a^k\cdot b^{((n+1)-k)}\right)
\end{aligned}$$

$$\begin{aligned}
&= \binom{n}{n} \cdot a^{n+1} \cdot b^{((n+1)-(n+1))} + \sum_{k=1}^{n} \left(\binom{n}{k-1} \cdot a^k \cdot b^{((n+1)-k)}\right) + \sum_{k=1}^{n} \left(\binom{n}{k} \cdot a^k \cdot b^{((n+1)-k)}\right) + \binom{0}{0} \cdot a^0 \cdot b^{((n+1)-0)} \\
&= a^{n+1} \cdot b^0 + \sum_{k=1}^{n} \left(\left(\binom{n}{k-1} + \binom{n}{k}\right) \cdot a^k \cdot b^{((n+1)-k)}\right) + a^0 \cdot b^{n+1} \\
&\underset{13.21}{=} \binom{n+1}{0} \cdot a^{n+1} \cdot b^{((n+1)-(n+1))} + \sum_{k=1}^{n} \left(\binom{n+1}{k} \cdot a^k \cdot b^{((n+1)-k)}\right) + \binom{n+1}{0} \cdot a^0 \cdot b^{n+1} \\
&= \sum_{k=0}^{n+1} \left(\binom{n+1}{k} \cdot a^k \cdot b^{((n+1)-k)}\right)
\end{aligned}$$

proving that $n+1 \in \mathcal{S}$ $\square$

Corollary 13.23. *Let $p: \mathbb{C} \rightarrow \mathbb{C}$ is a non constant polynomal and $z_0 \in \mathbb{C}$ then $p_{\rightarrow z_0}: \mathbb{C} \rightarrow \mathbb{C}$ defined by $p_{\rightarrow z_0}(z) = p(z + z_0)$ is a non constant polynomal of order n*

Proof. We prove this by induction so let $\mathcal{S} = \{n \in \mathbb{N}_0 \,|\, p$ is a polynomal with $\operatorname{ord}(p) \leqslant n$ then $p_{\rightarrow z_0}$ is a polynomal of order $n\}$ then we have

$1 \in \mathcal{S}$. If p is a polynomal of order 1 then $p(z) = a_1 \cdot z + a_0$ hence $p_{\rightarrow z_0}(z) = p(z + z_0) = a_1 \cdot (z + z_0) + a_0 = a_1 \cdot z + (a_1 \cdot z_0 + a_0)$ a polynomal of order 1 proving that $1 \in \mathcal{S}$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$. Let p be a polynomal with $\operatorname{ord}(p) \leqslant n+1$ then we have either

$\operatorname{ord}(p) \leqslant n$. then as $n \in \mathcal{S}$ we have $p_{\rightarrow z_0}$ is a polynomal with order n

$\operatorname{ord}(p) = n+1$. then there exists a $\{a_i\}_{i \in \{1,\ldots,n+1\}} \subseteq \mathbb{C}$ with $a_{n+1} \neq 0$ such that

$$\begin{aligned}
p(z) &= \sum_{i=0}^{n+1} a_i \cdot z^i \\
&= \sum_{i=1}^{n+1} a_i \cdot z^i + a_0 \\
&= \sum_{i=1}^{n+1} z \cdot (a_i \cdot z^{i-1}) + a_0 \\
&= z \cdot \sum_{i=1}^{n+1} a_i \cdot z^{i-1} + a_0
\end{aligned}$$

$$\begin{aligned}
&= z\cdot\sum_{i=0}^{n} a_{i+1}\cdot z^i + a_0\\
&= z\cdot\sum_{i=0}^{n} b_i\cdot z^i + a_0 \text{ where } \{b_i\}_{i\in\{0,\ldots,n\}} \text{ is defined by } b_i = a_{i+1}\\
&= z\cdot q(z) + a_0 \text{ where } q \text{ is a polynomal with } \operatorname{ord}(q) = n \text{ (as } b_n =\\
&\quad a_{n+1}\neq 0)
\end{aligned}$$

As $n\in\mathcal{S}$ we have that $q_{\rightarrow z_0}$ is a polynomal of order n so that there exists a $\{c_i\}_{i\in 0,\ldots,n}\subseteq\mathbb{C}$ such that $c_n\neq 0$ and $q(z+z_0) = q_{\rightarrow z_0}(z) = \sum_{i=0}^{n} c_i\cdot z^i$ hence we have

$$\begin{aligned}
p_{\rightarrow z_0}(z) &= (z+z_0)\cdot q(z+z_0) + a_0\\
&= (z+z_0)\cdot q_{\rightarrow z_0}(z) + a_0\\
&= (z+z_0)\cdot\sum_{i=0}^{n} c_i\cdot z^i + a_0\\
&= \sum_{i=0}^{n} c_i\cdot z^{i+1} + \sum_{i=0}^{n}(z_0\cdot c_i)\cdot z^i + a_0\\
&= \sum_{i=1}^{n+1} c_{i-1}\cdot z^i + \sum_{i=1}^{n}(z_0\cdot c_i)\cdot z^i + (z_0\cdot c_0 + a_0)\\
&= c_n\cdot z^{n+1} + \sum_{i=1}^{n}(c_{i-1} + z_0\cdot c_i)\cdot z^i + (z_0\cdot c_0 + a_0)\\
&= \sum_{i=0}^{n+1} d_i\cdot z^i \text{ where } \{d_i\}_{i\in\{0,\ldots,n+1\}} \text{ is defined by } d_i =\\
&\quad \begin{cases} c_n \text{ if } i = n+1\\ (c_{i-1} + z_0\cdot c_i) \text{ if } i\in\{1,\ldots,n\}\\ z_0\cdot c_0 + a_0\end{cases} \text{ so that } d_{n+1}\neq 0
\end{aligned}$$

which proved that $p_{\rightarrow z_0}$ is a polynomal of order $n+1$ proving that $n+1\in\mathcal{S}$ □

Lemma 13.24. *Let $p:\mathbb{C}\rightarrow\mathbb{C}$ be a non constant polynomal of order $n\in\mathbb{N}$ then there exists a $1\leqslant k\leqslant n$ and a polynomal q of order $n-k$ such that $p(z) = p(0) + z^k\cdot q(z)$ and $q(0)\neq 0$*

Proof. As p is a polynomal of order n there exists a $\{a_i\}_{i\in\{0,\ldots,n\}}\subseteq\mathbb{C}$ with $a_n\neq 0$ so that $p(z) = \sum_{i=0}^{n} a_i\cdot z^i$. Hence $p(0) = \sum_{i=0}^{n} a_i\cdot 0^i = a_0$ further let $k = \min\{i\in\{1,\ldots,n\}|a_i\neq 0\}$ (which is defined as $a_n\neq 0$) then $p(z) = \sum_{i=0}^{n} a_i\cdot z^i = \sum_{i=1}^{n} a_i\cdot z^i + a_0 = \sum_{i=k}^{n} a_i\cdot z^i + p(0) = \sum_{i=0}^{n-k} a_{i+k}\cdot z^{i+k} + p(0) = z^k\cdot\sum_{i=0}^{n-k} a_{i+k}\cdot z^i + p(0) = p(0) + q(z)$ where $q(z) = \sum_{i=0}^{n-k} b_i$ and $\{b_i\}_{i\in\{0,\ldots,n-k\}}$ is defined by $b_i = a_{i+k}$. Finally $q(0) = b_0 = a_k\neq 0$. □

Lemma 13.25. *Let* $n\in\mathbb{N}$ *and* $\{a_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{C}$ *then* $\overline{\sum_{i=1}^n a_i}=\sum_{i=1}^n \bar{a}_i$

Proof. This is easy proved by induction so let $S=\{n\in\mathbb{N}|\forall\{a_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{C}$ we have $\overline{\sum_{i=1}^n a_i}=\sum_{i=1}^n \bar{a}_i\}$ then we have

$1\in\mathcal{S}$. this follows from $\overline{\sum_{i=1}^1 a_i}=\bar{a}_1=\sum_{i=1}^1\bar{a}_i$ for $\{a_i\}_{i\in\{1,\ldots,1\}}\subseteq\mathbb{C}$

$n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$. we have for $\{a_i\}_{i\in\{1,\ldots,n+1\}}$ that

$$\begin{aligned}\overline{\sum_{i=1}^{n+1} a+i} &= \overline{a_{n+1}+\sum_{i=1}^n a_i}\\ &= \overline{a_{n+1}}+\overline{\sum_{i=1}^n a_i}\\ &\underset{n\in\mathcal{S}}{=} \overline{a_{n+1}}+\sum_{i=1}^n \bar{a}_i\\ &= \sum_{i=1}^{n+1}\bar{a}_i\end{aligned}$$

proving that $n+1\in\mathcal{S}$ □

The following lemma shows how we can split up a set $\{1,\ldots,2\cdot n\}$ in a set of even and odd numbers.

Lemma 13.26. *Let* $n\in\mathbb{N}$ *then* $\{0,\ldots,2\cdot n\}=\{2\cdot k|k\in\{0,\ldots,n\}\}\bigsqcup\{2\cdot k+1|k\in\{0,\ldots,n-1\}\}$ *and* $\{0,\ldots,2\cdot n-1\}=\{2\cdot k|k\in\{0,\ldots,n-1\}\}\bigsqcup\{2\cdot k+1|k\in\{0,\ldots,n-1\}\}$

Proof. We prove this by induction so let $\mathcal{S}=\{n\in\mathbb{N}|\{0,\ldots,2\cdot n\}=\{2\cdot k|k\in\{0,\ldots,n\}\}\bigcup\{2\cdot k+1|k\in\{1,\ldots,n-1\}\}\}$ then we have

$1\in\mathcal{S}$. then $\{0,1,2\}=\{0,2\}\bigcup\{1\}=\{2\cdot k|k\in\{0,1\}\}\bigcup\{2\cdot k+1|k\in\{0\}\}$ proving that $1\in\mathcal{S}$

$n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$.

$$\begin{aligned}\{1,\ldots,2\cdot(n+1)\} &= \{1,\ldots,2\cdot n\}\bigcup\{2\cdot n+1,2\cdot(n+1)\}\\ &\underset{n\in\mathcal{S}}{=} \{2\cdot k|k\in\{0,\ldots,n\}\}\bigcup\{2\cdot k+1|k\in\{0,\ldots,n-1\}\}\bigcup\\ &\quad\{2\cdot n+1\}\bigcup\{2\cdot(n+1)\}\\ &= \{2\cdot k|k\in\{0,\ldots,n+1\}\}\bigcup\{2\cdot k+1|k\in\{0,\ldots,n\}\}\\ &= \{2\cdot k|k\in\{0,\ldots,n+1\}\}\bigcup\{2\cdot k+1|k\in\{0,\ldots,(n+1)-1\}\}\end{aligned}$$

proving that $n+1\in\mathcal{S}$.

It remains to prove that $\{2\cdot k|k\in\{0,\ldots,n\}\}\bigcap\{2\cdot k+1|k\in\{0,\ldots,n-1\}\}=\emptyset$ so assume that $l\in\{2\cdot k|k\in\{0,\ldots,n\}\}\bigcap\{2\cdot k+1|k\in\{0,\ldots,n-1\}\}$ then there exists a $k\in\{0,\ldots,n\}$ and a $k'\in\{0,\ldots,n-1\}$ such that $2\cdot k=l=2\cdot k'+1$ which means that a even number is equal to a odd number which is not possible (see 6.56). Hence $\{2\cdot k|k\in\{0,\ldots,n\}\}\bigcap\{2\cdot k+1|k\in\{0,\ldots,n-1\}\}=\emptyset$.

For the second part let $\mathcal{S}=\{n\in\mathbb{N}|\{0,\ldots,2\cdot n-1\}=\{2\cdot k|k\in\{0,\ldots,n-1\}\}\bigcup\{2\cdot k+1|k\in\{0,\ldots,n-1\}\}\}$ then we have

$1\in\mathcal{S}$. then $\{0,1\}=\{0\}\bigcup\{1\}=\{2\cdot k|k\in\{0\}\}\bigcup\{2\cdot k+1|k\in\{0\}\}$ proving that $0\in\mathcal{S}$

$n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$.

$$\begin{aligned}\{1,\ldots,2\cdot(n+1)-1\} &= \{1,\ldots,2\cdot n+1\}\\ &= \{1,\ldots,2\cdot n-1\}\bigcup\{2\cdot n+1\}\bigcup\{2\cdot n\}\\ &\underset{n\in\mathcal{S}}{=} \{2\cdot k|k\in\{0,\ldots,n-1\}\}\bigcup\{2\cdot k+1|k\in\{0,\ldots,\\ &\quad n-1\}\}\bigcup\{2\cdot n+1\}\bigcup\{2\cdot n\}\\ &= \{2\cdot k|k\in\{0,\ldots,n\}\}\bigcup\{2\cdot k+1|k\in\{0,\ldots,n\}\}\\ &= \{2\cdot k|k\in\{0,\ldots,(n+1)-1\}\}\bigcup\{2\cdot k+1|k\in\\ &\quad \{0,\ldots,(n+1)-1\}\}\end{aligned}$$

proving that $n+1\in\mathcal{S}$

If $l\in\{2\cdot k|k\in\{0,\ldots,n-1\}\}\bigcap\{2\cdot k+1|k\in\{0,\ldots,n-1\}\}$ then there exists a $k\in\{0,\ldots,n-1\}$ and a $k'\in\{0,\ldots,n-1\}$ such that $2\cdot k=l=2\cdot k'+1$ which means that a even number is equal to a odd number which is not possible (see 6.56). Hence $\{2\cdot k|k\in\{0,\ldots,n-1\}\}\bigcap\{2\cdot k+1|k\in\{0,\ldots,n-1\}\}=\emptyset$ □

Lemma 13.27. *Let $n\in\mathbb{N}$ then if we define $A=\{(k,l)\in\{0,\ldots,n\}^2|k<l\}$ and $\forall m\in\{1,\ldots,2\cdot n-1\}$ $B_m=\{(k,l)\in A|k+l=m\}$ then $\bigsqcup_{m\in\{1,\ldots,n-1\}}B_m=A$*

Proof. As $\forall m\in\{1,\ldots,2\cdot n-1\}$ we have that $B_m\subseteq A$ (by definition) we must have that $\bigcup_{m\in\{1,\ldots,2\cdot n-1\}}B_m\subseteq A$. If $(k,l)\in A$ we have $0\leqslant k,l\leqslant n$ and as $k<l$ we have we have that $k+l<l+l\leqslant 2\cdot n$ so that $k+l\leqslant 2\cdot n-1$, further from $0\leqslant k<l$ we have $0<l\Rightarrow 0\leqslant k<k+l\Rightarrow 1\leqslant k+l$ hence $(k,l)\in B_{k+l}$ proving that $A\subseteq\bigcup_{m\in\{1,\ldots,n-1\}}B_m$. And thus taking in account the previous that

$$A=\bigcup_{m\in\{1,\ldots,2\cdot n-1\}}B_m$$

If now $m,m'\in\{1,\ldots,n-1\}$ with $m\neq m'$ then if $(k,l)\in B_m\bigcap B_{m'}$ we have that $m=k+l=m'$ contradicting that $m\neq m'$ so we have that

$$A=\bigsqcup_{m\in\{1,\ldots,2\cdot n-1\}}B_m$$ □

Lemma 13.28. *Let $n\in\mathbb{N}$ we have for $k=2\cdot n$ and $\zeta=\left(1+\frac{i}{k}\right)^2$ that $\mathrm{Re}(\zeta^k)<0<\mathrm{Img}(\zeta^k)$*

Proof. We have

$$\begin{aligned}\left(1+\frac{i}{k}\right)^{2\cdot k} &= \left(\frac{i}{k}+1\right)^{2\cdot k}\\ &= \sum_{l=0}^{2\cdot k}\left(\binom{2\cdot k}{l}\left(\frac{i}{k}\right)^{l}\cdot 1^{2\cdot k-l}\right)\\ &= \sum_{l\in\{0,\ldots,2\cdot k\}}\left(\binom{2\cdot k}{l}\left(\frac{i}{k}\right)^{l}\right)\\ &\underset{13.26}{=} \sum_{l\in\{2\cdot j|j\in\{0,\ldots,k\}\}\bigsqcup\{2\cdot j+1|j\in\{0,\ldots,k-1\}\}}\left(\binom{2\cdot k}{l}\left(\frac{i}{k}\right)^{l}\right)\\ &= \left(\sum_{l\in\{2\cdot j|j\in\{0,\ldots,k\}\}}\left(\binom{2\cdot k}{l}\left(\frac{i}{k}\right)^{l}\right)\right)+\\ &\quad\left(\sum_{l\in\{2\cdot j+1|j\in\{0,\ldots,k-1\}\}}\left(\binom{2\cdot k}{l}\left(\frac{i}{k}\right)^{l}\right)\right)\\ &= \left(\sum_{j\in\{0,\ldots,k\}}\left(\binom{2\cdot k}{2\cdot j}\left(\frac{i}{k}\right)^{2\cdot j}\right)\right)+\\ &\quad\left(\sum_{j\in\{0,\ldots,k-1\}}\left(\binom{2\cdot k}{2\cdot j+1}\left(\frac{i}{k}\right)^{2\cdot j+1}\right)\right)\end{aligned}$$

Further as $k=2\cdot n$

$$\begin{aligned}\sum_{j\in\{0,\ldots,k\}}\left(\binom{2\cdot k}{2\cdot j}\left(\frac{i}{k}\right)^{2\cdot j}\right) &= \sum_{j\in\{0,\ldots,2\cdot n\}}\left(\binom{2\cdot k}{2\cdot j}\left(\frac{i}{k}\right)^{2\cdot j}\right)\\ &\underset{13.26}{=} \sum_{j\in\{2\cdot l|l\in\{0,\ldots,n\}\}\bigsqcup\{2\cdot l+1|l\in\{0,\ldots,n-1\}\}}\left(\binom{2\cdot k}{2\cdot j}\left(\frac{i}{k}\right)^{2\cdot j}\right)\\ &= \left(\sum_{j\in\{2\cdot l|l\in\{0,\ldots,n\}\}}\left(\binom{2\cdot k}{2\cdot j}\left(\frac{i}{k}\right)^{2\cdot j}\right)\right)+\\ &\quad\left(\sum_{j\in\{2\cdot l+1|l\in\{0,\ldots,n-1\}\}}\left(\binom{2\cdot k}{2\cdot j}\left(\frac{i}{k}\right)^{2\cdot j}\right)\right)\\ &= \left(\sum_{l\in\{0,\ldots,n\}}\left(\binom{2\cdot k}{2\cdot 2\cdot l}\left(\frac{i}{k}\right)^{2\cdot 2\cdot l}\right)\right)+\\ &\quad\left(\sum_{l\in\{0,\ldots,n-1\}}\left(\binom{2\cdot k}{2\cdot(2\cdot l+1)}\left(\frac{i}{k}\right)^{2\cdot(2\cdot l+1)}\right)\right)\end{aligned}$$

$$
\begin{aligned}
&= \left(\sum_{l\in\{0,\ldots,n\}} \left(\binom{2\cdot k}{4\cdot l}\left(\frac{i}{k}\right)^{4\cdot l}\right)\right) + \\
&\quad \left(\sum_{l\in\{0,\ldots,n-1\}} \left(\binom{2\cdot k}{4\cdot l+2}\left(\frac{i}{k}\right)^{4\cdot l+2}\right)\right) \\
&\underset{9.22}{=} \left(\sum_{l\in\{0,\ldots,n\}} \left(\binom{2\cdot k}{4\cdot l} \cdot \frac{1}{k^{4\cdot l}}\right)\right) - \\
&\quad \left(\sum_{l\in\{0,\ldots,n-1\}} \left(\binom{2\cdot k}{4\cdot l+2}\cdot\frac{1}{k^{4\cdot l+2}}\right)\right) \\
&= \binom{2\cdot k}{4\cdot 0}\frac{1}{k^{4\cdot 0}} + \binom{2\cdot k}{4} \cdot \frac{1}{k^4} - \binom{2\cdot k}{4\cdot 0+2} \cdot \\
&\quad \frac{1}{k^{4\cdot 0+2}} + \sum_{l\in\{2,\ldots,n\}} \left(\binom{2\cdot k}{4\cdot l} \cdot \frac{1}{k^{4\cdot l}}\right) - \\
&\quad \sum_{l\in\{1,\ldots,n-1\}} \left(\binom{2\cdot k}{4\cdot l+2}\cdot\frac{1}{k^{4\cdot l+2}}\right) \\
&= 1 - \binom{2\cdot k}{2} \cdot \frac{1}{k^2} + \binom{2\cdot k}{4} \cdot \\
&\quad \frac{1}{k^4} + \sum_{l\in\{2,\ldots,n\}} \left(\binom{2\cdot k}{4\cdot l} \cdot \frac{1}{k^{4\cdot l}}\right) - \\
&\quad \sum_{l\in\{1,\ldots,n-1\}} \left(\binom{2\cdot k}{4\cdot l+2}\cdot\frac{1}{k^{4\cdot l+2}}\right) \\
&= 1 - \binom{2\cdot k}{2} \cdot \frac{1}{k^2} + \binom{2\cdot k}{4} \cdot \frac{1}{k^4} + \\
&\quad \sum_{l\in\{1,\ldots,n-1\}} \left(\binom{2\cdot k}{4\cdot(l+1)} \cdot \frac{1}{k^{4\cdot(l+1)}}\right) - \\
&\quad \sum_{l\in\{1,\ldots,n-1\}} \left(\binom{2\cdot k}{2\cdot(2\cdot l+1)}\cdot\frac{1}{k^{2\cdot(2\cdot l+1)}}\right) \\
&= 1 - \binom{2\cdot k}{2} \cdot \frac{1}{k^2} + \binom{2\cdot k}{4} \cdot \frac{1}{k^4} + \\
&\quad \sum_{l\in\{1,\ldots,n-1\}} \left(\binom{2\cdot k}{2\cdot(2\cdot l+1)+2} \cdot \right. \\
&\quad \left.\frac{1}{k^{2\cdot(2\cdot l+1)+2}}\right) - \sum_{l\in\{2\cdot l+1|l\in\{1,\ldots,n-1\}\}} \left(\binom{2\cdot k}{2\cdot j}\cdot \right. \\
&\quad \left.\frac{1}{k^{2\cdot j}}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= 1 - \binom{2\cdot k}{2} \cdot \frac{1}{k^2} + \binom{2\cdot k}{4} \cdot \frac{1}{k^4} + \\
&\quad \sum_{j\in\{2\cdot l+1 \mid l\in\{1,\dots,n-1\}\}} \left(\binom{2\cdot k}{2\cdot j+2}\cdot\frac{1}{k^{2\cdot j+2}}\right) - \\
&\quad \sum_{l\in\{2\cdot l+1 \mid l\in\{1,\dots,n-1\}\}} \left(\binom{2\cdot k}{2\cdot j}\cdot\frac{1}{k^{2\cdot j}}\right) \\
&= 1 - \binom{2\cdot k}{2} \cdot \frac{1}{k^2} + \binom{2\cdot k}{4} \cdot \frac{1}{k^4} + \\
&\quad \sum_{j\in\{2\cdot l+1 \mid l\in\{1,\dots,n-1\}\}} \left(\binom{2\cdot k}{2\cdot j+2}\cdot\frac{1}{k^{2\cdot j+2}} - \right. \\
&\quad \left.\binom{2\cdot k}{2\cdot j}\cdot\frac{1}{k^{2\cdot j}}\right)
\end{aligned}
$$

$$
\begin{aligned}
\sum_{j\in\{0,\dots,k-1\}} \left(\binom{2\cdot k}{2\cdot j+1}\left(\frac{i}{k}\right)^{2\cdot j+1}\right) &= \sum_{j\in\{0,\dots,2\cdot n-1\}} \left(\binom{2\cdot k}{2\cdot j+1}\left(\frac{i}{k}\right)^{2\cdot j+1}\right) \\
&= \sum_{j\in\{2\cdot l \mid l\in\{0,\dots,n-1\}\}} \left(\binom{2\cdot k}{2\cdot j+1}\left(\frac{i}{k}\right)^{2\cdot j+1}\right) + \\
&\quad \sum_{j\in\{2\cdot l+1 \mid l\in\{0,\dots,n-1\}\}} \left(\binom{2\cdot k}{2\cdot j+1}\left(\frac{i}{k}\right)^{2\cdot j+1}\right) \\
&= \sum_{l\in\{0,\dots,n-1\}} \left(\binom{2\cdot k}{2\cdot(2\cdot l)+1}\left(\frac{i}{k}\right)^{2\cdot(2\cdot l)+1}\right) + \\
&\quad \sum_{l\in\{0,\dots,n-1\}} \left(\binom{2\cdot k}{2\cdot(2\cdot l+1)+1}\left(\frac{i}{k}\right)^{2\cdot(2\cdot l+1)+1}\right) \\
&= \sum_{l\in\{0,\dots,n-1\}} \left(\binom{2\cdot k}{4\cdot l+1}\left(\frac{i}{k}\right)^{4\cdot l+1}\right) + \\
&\quad \sum_{l\in\{0,\dots,n-1\}} \left(\binom{2\cdot k}{4\cdot l+3}\left(\frac{i}{k}\right)^{4\cdot l+3}\right) \\
&\underset{9.22}{=} \sum_{l\in\{0,\dots,n-1\}} \left(\binom{2\cdot k}{4\cdot l+1}\cdot\frac{i}{k^{4\cdot l+1}}\right) + \\
&\quad \sum_{l\in\{0,\dots,n-1\}} \left(\binom{2\cdot k}{4\cdot l+3}\cdot\frac{-i}{k^{4\cdot l+3}}\right) \\
&= i\cdot \sum_{l\in\{0,\dots,n-1\}} \left(\binom{2\cdot k}{4\cdot l+1}\cdot\frac{1}{k^{4\cdot l+1}} - \right. \\
&\quad \left.\binom{2\cdot k}{4\cdot l+3}\cdot\frac{1}{k^{4\cdot l+3}}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= i\cdot\sum_{l\in\{0,\ldots,n-1\}}\left(\binom{2\cdot k}{2\cdot(2\cdot l+1)-1}\cdot\frac{1}{k^{2\cdot(2\cdot l+1)-1}}-\binom{2\cdot k}{2\cdot(2\cdot l+1)+1}\cdot\frac{1}{k^{2\cdot(2\cdot l+1)+1}}\right)\\
&= i\cdot\sum_{j\in\{2\cdot l+1|l\in\{0,\ldots,n-1\}\}}\left(\binom{2\cdot k}{2\cdot j-1}\cdot\frac{1}{k^{2\cdot j-1}}-\binom{2\cdot k}{2\cdot j+1}\cdot\frac{1}{k^{2\cdot j+1}}\right)
\end{aligned}
$$

To summarize we have that

$$
\mathrm{Re}\left(\left(1+\frac{i}{k}\right)^{2\cdot k}\right)=1-\binom{2\cdot k}{2}\cdot\frac{1}{k^2}+\binom{2\cdot k}{4}\cdot\frac{1}{k^4}+\sum_{j\in\{2\cdot l+1|l\in\{1,\ldots,n-1\}\}}\left(\binom{2\cdot k}{2\cdot j+2}\cdot\frac{1}{k^{2\cdot j+2}}-\binom{2\cdot k}{2\cdot j}\cdot\frac{1}{k^{2\cdot j}}\right) \tag{13.6}
$$

$$
\mathrm{Img}\left(\left(1+\frac{i}{k}\right)^{2\cdot k}\right)=\sum_{j\in\{2\cdot l+1|l\in\{0,\ldots,n-1\}\}}\left(\binom{2\cdot k}{2\cdot j-1}\cdot\frac{1}{k^{2\cdot j-1}}-\binom{2\cdot k}{2\cdot j+1}\cdot\frac{1}{k^{2\cdot j+1}}\right) \tag{13.7}
$$

Further we have that as $2\leqslant k\Rightarrow 4\leqslant 2\cdot k$

$$
\begin{aligned}
\binom{2\cdot k}{2} &= \frac{(2\cdot k)!}{2!\cdot(2\cdot k-2)!}\\
&= \frac{(2\cdot k)\cdot(2\cdot k-1)\cdot(2\cdot k-2)!}{2\cdot(2\cdot k-2)!}\\
&= k\cdot(2\cdot k-1)\\
&= 2\cdot k^2-k\\
\binom{2\cdot k}{4} &= \frac{(2\cdot k)!}{4!\cdot(2\cdot k-4)!}\\
&= \frac{(2\cdot k)!\cdot(2\cdot k-3)\cdot(2\cdot k-2)}{4\cdot 3\cdot 2!\cdot(2\cdot k-4)!\cdot(2\cdot k-3)\cdot(2\cdot k-2)}\\
&= \frac{(2\cdot k)!\cdot(2\cdot k-3)\cdot(2\cdot k-2)}{4\cdot 3\cdot 2!\cdot(2\cdot k-2)!}\\
&= \frac{(2\cdot k-3)\cdot(2\cdot k-2)}{12}\cdot\binom{2\cdot k}{2}\\
&= \binom{2\cdot k}{2}\cdot\frac{(2\cdot k-3)\cdot(k-1)}{6}
\end{aligned}
$$

$$\begin{aligned}
&= \binom{2\cdot k}{2}\cdot\frac{2\cdot k^2-2\cdot k-3\cdot k+3}{6}\\
&= \binom{2\cdot k}{2}\cdot\frac{2\cdot k^2-5\cdot k+3}{6}
\end{aligned}$$

Hence

$$\begin{aligned}
1-\binom{2\cdot k}{2}\cdot\frac{1}{k^2}+\binom{2\cdot k}{4}\cdot\frac{1}{k^4} &= 1-\binom{2\cdot k}{2}\cdot\left[\frac{1}{k^2}-\frac{2\cdot k^2-5\cdot k+3}{6\cdot k^4}\right]\\
&= 1-\binom{2\cdot k}{2}\cdot\frac{1}{k^2}\cdot\left[1-\frac{2\cdot k^2-5\cdot k+3}{6\cdot k^2}\right]\\
&= 1-\frac{2\cdot k^2-k}{k^2}\cdot\left[1-\frac{1}{3}+\frac{5\cdot k-3}{6\cdot k^2}\right]\\
&= 1-\left(2-\frac{1}{k}\right)\cdot\left(\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}\right)
\end{aligned}$$

As $0<2\leqslant k$ we have $\frac{1}{k}\leqslant\frac{1}{2}\Rightarrow-\frac{1}{2}\leqslant-\frac{1}{k}$ hence $\left(2-\frac{1}{k}\right)\geqslant\left(2-\frac{1}{2}\right)=\frac{3}{2}\Rightarrow-\frac{3}{2}\geqslant\left(2-\frac{1}{k}\right)$ further as $2\leqslant k\Rightarrow 10\leqslant 5\cdot k\Rightarrow 0<7\leqslant 5\cdot k-3$ hence $0\leqslant\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}$ so $-\left(2-\frac{1}{k}\right)\cdot\left(\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}\right)\leqslant-\frac{3}{2}\cdot\left(\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}\right)$ form which we have

$$\begin{aligned}
1-\left(2-\frac{1}{k}\right)\cdot\left(\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}\right) &\leqslant 1-\frac{3}{2}\cdot\left(\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}\right)\\
&= 1-1-\frac{3}{2}\cdot\frac{5\cdot k-3}{6\cdot k^2}\\
&= \cdot\frac{5\cdot k-3}{4\cdot k^2}\\
&< 0
\end{aligned}$$

So we have

$$1-\binom{2\cdot k}{2}\cdot\frac{1}{k^2}+\binom{2\cdot k}{4}\cdot\frac{1}{k^4}<0 \tag{13.8}$$

Further for $j\in\{2\cdot l+1|l\in\{1,\ldots,n-1\}\}$ we have $1\leqslant j\leqslant 2\cdot(n-1)+1=2\cdot n-1=k-1<k$ hence

$$\begin{aligned}
\binom{2\cdot k}{2\cdot j+2}\cdot\frac{1}{k^{2\cdot j+2}}-\binom{2\cdot k}{2\cdot j}\cdot\frac{1}{k^{2\cdot j}} &= \frac{(2\cdot k)!}{(2\cdot j+2)!\cdot(2\cdot k-2\cdot j-2)!\cdot k^{2\cdot j+2}} -\\
&\quad \frac{(2\cdot k)!}{(2\cdot j)!\cdot(2\cdot k-2\cdot j)!\cdot k^{2\cdot j}}\\
&= \frac{(2\cdot k)!}{k^{2\cdot j}}\cdot\left(\frac{1}{(2\cdot j+2)!\cdot(2\cdot k-2\cdot j-2)!\cdot k^2}-\right.\\
&\quad\left.\frac{1}{(2\cdot j)!\cdot(2\cdot k-2\cdot j)!}\right)
\end{aligned}$$

$$\begin{aligned}
&= \frac{(2\cdot k)!}{k^{2\cdot j}}\cdot\left(\frac{1}{(2\cdot j+2)\cdot(2\cdot j+1)\cdot(2\cdot j)!\cdot(2\cdot k-2\cdot j-2)!\cdot k^2}-\right.\\
&\quad\left.\frac{1}{(2\cdot j)!\cdot(2\cdot k-2\cdot j)!}\right)\\
&= \frac{(2\cdot k)!}{(2\cdot j)!\cdot k^{2\cdot j}}\cdot\left(\frac{1}{(2\cdot j+2)\cdot(2\cdot j+1)\cdot(2\cdot k-2\cdot j-2)!\cdot k^2}-\right.\\
&\quad\left.\frac{1}{(2\cdot k-2\cdot j)!}\right)\\
&= \frac{(2\cdot k)!}{(2\cdot j)!\cdot k^{2\cdot j}}\cdot\left(\frac{1}{(2\cdot j+2)\cdot(2\cdot j+1)\cdot(2\cdot k-2\cdot j-2)!\cdot k^2}-\right.\\
&\quad\left.\frac{1}{(2\cdot k-2\cdot j)\cdot(2\cdot k-2\cdot j-1)\cdot(2\cdot k-2\cdot j-2)!}\right)\\
&= \frac{(2\cdot k)!}{(2\cdot j)!\cdot(2\cdot k-2\cdot j-2)!\cdot k^{2\cdot j}}\cdot\\
&\quad\left(\frac{1}{(2\cdot j+2)\cdot(2\cdot j+1)\cdot k^2}-\right.\\
&\quad\left.\frac{1}{(2\cdot k-2\cdot j)\cdot(2\cdot k-2\cdot j-1)}\right)
\end{aligned}$$

proving that

$$\binom{2\cdot k}{2\cdot j+2}\cdot\frac{1}{k^{2\cdot j+2}}-\binom{2\cdot k}{2\cdot j}\cdot\frac{1}{k^{2\cdot j}} = \frac{(2\cdot k)!}{(2\cdot j)!\cdot(2\cdot k-2\cdot j-2)!\cdot k^{2\cdot j}}\cdot \left(\frac{1}{(2\cdot j+2)\cdot(2\cdot j+1)\cdot k^2}-\frac{1}{(2\cdot k-2\cdot j)\cdot(2\cdot k-2\cdot j-1)}\right) \tag{13.9}$$

Now as $0\leqslant 1<j<k$ we have

$$0<(2\cdot j+2)\cdot(2\cdot j+1)\cdot k^2,(2\cdot k-2\cdot j)\cdot(2\cdot k-2\cdot j-1) \tag{13.10}$$

and

$$\begin{aligned}
&(2\cdot j+2)\cdot(2\cdot j+1)\cdot k^2-(2\cdot k-2\cdot j)\cdot\\
&(2\cdot k-2\cdot j-1) = 2\cdot[(j+1)\cdot(2\cdot j+1)\cdot k^2-(k-1)\cdot(2\cdot k-2\cdot j-1)]\\
&= 2\cdot[2\cdot j^2\cdot k^2+j\cdot k^2+2\cdot j\cdot k^2+k^2-[2\cdot k^2-2\cdot k\cdot j-k-2\cdot k+2\cdot j+1]]\\
&= 2\cdot[2\cdot j^2\cdot k^2+3\cdot j\cdot k^2+k^2-2\cdot k^2+2\cdot k\cdot j+k+2\cdot k-2\cdot j-1]\\
&= 2\cdot[2\cdot j^2\cdot k^2+3\cdot j\cdot k^2-k^2+2\cdot k\cdot j+k+2\cdot k-2\cdot j-1]\\
&> 2\cdot[2\cdot j^2\cdot k^2+3\cdot k^2-k^2+2\cdot k\cdot j+k+2\cdot k-2\cdot j-1]
\end{aligned}$$

$$\begin{aligned} &= 2\cdot[2\cdot j^2\cdot k^2+2\cdot k^2+2\cdot k\cdot j+(k-1)+\\ &\quad 2\cdot(k-j)]\\ &> 0\end{aligned}$$

hence using the above, 9.41 and 13.10 we have that $\left(\frac{1}{(2\cdot j+2)\cdot(2\cdot j+1)\cdot k^2}-\frac{1}{(2\cdot k-2\cdot j)\cdot(2\cdot k-2\cdot j-1)}\right)<0$ so that by 13.9 we have

$$\binom{2\cdot k}{2\cdot j+2}\cdot\frac{1}{k^{2\cdot j+2}}-\binom{2\cdot k}{2\cdot j}\cdot\frac{1}{k^{2\cdot j}}<0 \tag{13.11}$$

Combining then 13.8 and 13.11 on 13.6 proves that

$$\mathrm{Re}\left(\left(1+\frac{i}{k}\right)^{2\cdot k}\right)<0 \tag{13.12}$$

Now to extmiate the imaginair part note that

$$\begin{aligned}
&\binom{2\cdot k}{2\cdot j-1}\cdot\frac{1}{k^{2\cdot j-1}}-\binom{2\cdot k}{2\cdot j+1}\cdot\frac{1}{k^{2\cdot j+1}}\\
&= \frac{1}{k^{2\cdot j-1}}\cdot\left(\binom{2\cdot k}{2\cdot j-1}-\binom{2\cdot k}{2\cdot j+1}\cdot\frac{1}{k^2}\right)\\
&= \frac{1}{k^{2\cdot j-1}}\cdot\left(\frac{(2\cdot k)!}{(2\cdot j-1)!\cdot(2\cdot k-2\cdot j+1)!}-\frac{(2\cdot k)!}{(2\cdot j+1)!(2\cdot k-2\cdot j-1)}\cdot\frac{1}{k^2}\right)\\
&= \frac{(2\cdot k)!}{k^{2\cdot j-1}}\cdot\left(\frac{1}{(2\cdot j-1)!\cdot(2\cdot k-2\cdot j+1)!}-\frac{1}{(2\cdot j+1)!(2\cdot k-2\cdot j-1)!\cdot k^2}\right)\\
&= \frac{(2\cdot k)!}{k^{2\cdot j-1}}\cdot\left(\frac{1}{(2\cdot j-1)!\cdot(2\cdot k-2\cdot j+1)!}-\frac{1}{(2\cdot j+1)\cdot(2\cdot j)\cdot(2\cdot j-1)!\cdot(2\cdot k-2\cdot j-1)!\cdot k^2}\right)\\
&= \frac{(2\cdot k)!}{k^{2\cdot j-1}\cdot(2\cdot j-1)!}\cdot\left(\frac{1}{(2\cdot k-2\cdot j+1)!}-\frac{1}{(2\cdot j+1)\cdot(2\cdot j)\cdot(2\cdot k-2\cdot j-1)!\cdot k^2}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{(2\cdot k)!}{k^{2\cdot j-1}\cdot(2\cdot j-1)!}\cdot \\
&\quad \left(\frac{1}{(2\cdot k-2\cdot j+1)\cdot(2\cdot k-2\cdot j)\cdot(2\cdot k-2\cdot j-1)!}-\right. \\
&\quad \left.\frac{1}{(2\cdot j+1)\cdot(2\cdot j)\cdot(2\cdot k-2\cdot j-1)!\cdot k^2}\right) \\
&= \frac{(2\cdot k)!}{k^{2\cdot j-1}\cdot(2\cdot j-1)!\cdot(2\cdot k-2\cdot j-1)!}\cdot \\
&\quad \left(\frac{1}{(2\cdot k-2\cdot j+1)\cdot(2\cdot k-2\cdot j)}-\right. \\
&\quad \left.\frac{1}{(2\cdot j+1)\cdot(2\cdot j)\cdot k^2}\right)
\end{aligned}
$$

Proving that

$$
\binom{2\cdot k}{2\cdot j-1}\cdot\frac{1}{k^{2\cdot j-1}}-\binom{2\cdot k}{2\cdot j+1}\cdot\frac{1}{k^{2\cdot j+1}}=\left(\frac{1}{(2\cdot k-2\cdot j+1)\cdot(2\cdot k-2\cdot j)}-\frac{1}{(2\cdot j+1)\cdot(2\cdot j)\cdot k^2}\right) \tag{13.13}
$$

Now we have using $1<j<k$ that $2\cdot k-2\cdot j>0$ hence

$$
0<(2\cdot k-2\cdot j+1)\cdot(2\cdot k-2\cdot j),(2\cdot j+1)\cdot(2\cdot j)\cdot k^2 \tag{13.14}
$$

further taking in account that form $0<1\leqslant j$ we have by 9.41 that $j^2\geqslant 1$

$$
\begin{aligned}
(2\cdot j+1)\cdot(2\cdot j)\cdot k^2-(2\cdot k-2\cdot j+1)\cdot(2\cdot k-2\cdot j) &= 2\cdot[(2\cdot j+1)\cdot j\cdot k^2-(2\cdot k-2\cdot j+1)\cdot(k-j)] \\
&= 2\cdot[2\cdot j^2\cdot k^2+j\cdot k^2-[2\cdot k^2-2\cdot k\cdot j+k-2\cdot k\cdot j-2\cdot j^2-j]] \\
&= 2\cdot[2\cdot j^2\cdot k^2+j\cdot k^2-2\cdot k^2+2\cdot k\cdot j-k+2\cdot k\cdot j+2\cdot j^2+j] \\
&= 2\cdot[2\cdot j^2\cdot k^2+j\cdot k^2-2\cdot k^2+4\cdot k\cdot j-k+2\cdot j^2+j] \\
&\geqslant 2\cdot[2\cdot k^2+j\cdot k^2-2\cdot k^2+4\cdot k-k+2\cdot j^2+j] \\
&= 2\cdot[j\cdot k^2 3\cdot k+2\cdot j^2+j] \\
&> 0
\end{aligned}
$$

Hence using the above together with 13.14 gives $\frac{1}{(2\cdot k-2\cdot j+1)\cdot(2\cdot k-2\cdot j)}-\frac{1}{(2\cdot j+1)\cdot(2\cdot j)\cdot k^2}=\frac{(2\cdot j+1)\cdot(2\cdot j)\cdot k^2-(2\cdot k-2\cdot j+1)\cdot(2\cdot k-2\cdot j)}{(2\cdot k-2\cdot j+1)\cdot(2\cdot k-2\cdot j)\cdot(2\cdot j+1)\cdot(2\cdot j)\cdot k^2}>0$ which using 13.13 proves that $\binom{2\cdot k}{2\cdot j-1}\cdot\frac{1}{k^{2\cdot j-1}}-\binom{2\cdot k}{2\cdot j+1}\cdot\frac{1}{k^{2\cdot j+1}}>0$. Substituting this in 13.7

gives

$$\mathrm{Img}\left(\left(1+\frac{i}{k}\right)^{2\cdot k}\right)>0 \tag{13.15}$$

The theorem is then proved by 13.12 and 13.15 □

13.1.4 Proof of the fundamental theorem of algebra

After all this work we are ready for the fundamental theorem of algebra.

Theorem 13.29. *Let $n\in\mathbb{N}$, $\{a_i\}_{i\in\{0,\ldots,n\}}\subseteq\mathbb{C}$ with $a_n\neq 0$ and $p:\mathbb{C}\rightarrow\mathbb{C}$ defined by $p(z)=\sum_{i=1}^n a_i\cdot z^i$ (p is a non constant polynomial) then there exists a $z_0\in\mathbb{C}$ such that $p(z_0)=0$ (in other words every non constant polynomal has a zero point)*

Proof. First define $p\cdot\bar{p}:\mathbb{C}\rightarrow\mathbb{R}$ by $(p\cdot\bar{p})(z)=p(z)\cdot\overline{p(z)}\geqslant 0$ now given $z\in\mathbb{C}$ we have

$$\begin{aligned}
(p\cdot\bar{p})(z) &= p(z)\cdot\overline{p(z)}\\
&\underset{13.25}{=} \left(\sum_{i\in\{0,\ldots,n\}} a_i\cdot z^i\right)\cdot\overline{\left(\sum_{i\in\{0,\ldots,n\}} a_i\cdot z^i\right)}\\
&= \left(\sum_{i\in\{0,\ldots,n\}} a_i\cdot z^i\right)\cdot\left(\sum_{i\in\{0,\ldots,n\}} \overline{a_i\cdot z^i}\right)\\
&= \left(\sum_{i\in\{0,\ldots,n\}} a_i\cdot z^i\right)\cdot\left(\sum_{i\in\{0,\ldots,n\}} \bar{a_i}\cdot \bar{z^i}\right)\\
&\underset{10.61}{=} \sum_{(i,j)\in\{0,\ldots,n\}^2} a_i\cdot z^i\cdot\bar{a_j}\cdot\bar{z^j}\\
&= \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k=l\}\bigsqcup\{(k,l)\in\{0,\ldots,n\}^2|k<l\}\bigsqcup\{(k,l)\in\{0,\ldots,n\}^2|l<k\}} a_i\cdot\\
&\quad z^i\cdot\bar{a_j}\cdot\bar{z^j}\\
&= \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k=l\}} a_i\cdot z^i\cdot\bar{a_j}\cdot\bar{z^j}+\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} a_i\cdot\\
&\quad z^i\cdot\bar{a_j}\cdot\bar{z^j}+\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|l<k\}} a_i\cdot z^i\cdot\bar{a_j}\cdot\bar{z^j}\\
&= \sum_{i\in\{0,\ldots,n\}} a_i\cdot z^i\cdot\bar{a_i}\cdot\bar{z^i}+\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} a_i\cdot z^i\cdot\bar{a_j}\cdot\bar{z^j}+\\
&\quad \sum_{(j,i)\in\{(k,l)\in\{0,\ldots,n\}^2|l<k\}} a_j\cdot z^j\cdot\bar{a_i}\cdot\bar{z^i}\\
&= \sum_{i\in\{0,\ldots,n\}} a_i\cdot z^i\cdot\bar{a_i}\cdot\bar{z^i}+\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} a_i\cdot z^i\cdot\bar{a_j}\cdot\bar{z^j}+\\
&\quad \sum_{(i,j)\in\{(l,k)\in\{0,\ldots,n\}^2|k<l\}} a_j\cdot z^j\cdot\bar{a_i}\cdot\bar{z^i}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i\in\{0,\ldots,n\}} a_i\cdot z^i\cdot \bar{a_i}\cdot \bar{z^i} + \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} (a_i\cdot z^i\cdot \bar{a_j}\cdot \bar{z^j} + a_j\cdot z^j\cdot \bar{a_i}\cdot \bar{z^i})\\
&= \sum_{i\in\{0,\ldots,n\}} a_i\cdot z^i\cdot \bar{a_i}\cdot \bar{z^i} + \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} (a_i\cdot z^i\cdot \bar{a_j}\cdot \bar{z^j} + a_j\cdot z^j\cdot \bar{a_i}\cdot \bar{z^i})\\
&= \sum_{i\in\{0,\ldots,n\}} a_i\cdot z^i\cdot \bar{a_i}\cdot \bar{z^i} + \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} (a_i\cdot z^i\cdot \bar{a_j}\cdot \bar{z^j} + \overline{a_i\cdot z^i\cdot \bar{a_j}\cdot \bar{z^j}})\\
&= \sum_{i\in\{0,\ldots,n\}} a_i\cdot z^i\cdot \bar{a_i}\cdot \bar{z^i} + 2\cdot \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \mathrm{Re}(a_i\cdot z^i\cdot \bar{a_j}\cdot \bar{z^j})\\
&= \sum_{i\in\{0,\ldots,n\}} |a_i|^2\cdot |z^i|^2 + 2\cdot \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \mathrm{Re}(a_i\cdot z^i\cdot \bar{a_j}\cdot \bar{z^j})
\end{aligned}
$$

So we have that

$$
(p\cdot\bar{p})(z) = \sum_{i\in\{0,\ldots,n\}} |a_i|^2\cdot |z^i|^2 + 2\cdot \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \mathrm{Re}(a_i\cdot z^i\cdot \bar{a_j}\cdot \bar{z^j}) \tag{13.16}
$$

Using the above we have

$$
\begin{aligned}
(p\cdot\bar{p})(z) &= p(z)\cdot\overline{p(z)}\\
&\underset{0\leqslant p(z)\cdot\overline{p(z)}\in\mathbb{R}}{=} |p(z)\cdot\overline{p(z)}|\\
&\underset{p(z)\cdot\overline{p(z)}\in\mathbb{R} \text{ and } 13.15}{=} \|p(z)\cdot\overline{p(z)}\|\\
&= \left\| \sum_{i\in\{0,\ldots,n\}} |a_i|^i \cdot |z^i|^2 + 2\cdot \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \mathrm{Re}(a_i\cdot z^i\cdot \bar{a_j}\cdot \bar{z^j})\right\|\\
&\geqslant_{12.70} \left| \left\| \sum_{i\in\{0,\ldots,n\}} |a_i|^{2\cdot i} \cdot |z^i|^2 \right\| - \left\| 2\cdot \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \mathrm{Re}(a_i\cdot z^i\cdot \bar{a_j}\cdot \bar{z^j})\right\| \right|\\
&\geqslant \left\| \sum_{i\in\{0,\ldots,n\}} |a_i|^{2\cdot i} \cdot |z^i|^2 \right\| - 2\cdot \left\| \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \mathrm{Re}(a_i\cdot z^i\cdot \bar{a_j}\cdot \bar{z^j})\right\|
\end{aligned}
$$

$$= \sum_{i\in\{0,\ldots,n\}} |a_i|^{2\cdot i} \cdot |z^i|^2 - 2 \cdot \left\| \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2 | k<l\}} \mathrm{Re}(a_i \cdot z^i \cdot \bar{a_j} \cdot \bar{z^j}) \right\| \text{ as } [0 \leqslant \sum_{i\in\{0,\ldots,n\}} |a_i|^i \cdot |z|^i \in \mathbb{R} \text{ and } 13.15\ (1)]$$

$$\geqslant |a_n|^{2\cdot n} \cdot |z^n|^2 - 2 \cdot \left\| \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2 | k<l\}} \mathrm{Re}(a_i \cdot z^i \cdot \bar{a_j} \cdot \bar{z^j}) \right\|]$$

$$\geqslant_{13.15\ (3)} |a_n|^{2\cdot n} \cdot |z^n|^2 - 2 \cdot \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2 | k<l\}} \left\| a_i \cdot \bar{a_j} \right\| \cdot \|z^i\| \cdot \|\bar{z^j}\|$$

$$\geqslant_{13.15\ (5)} |a_n|^{2\cdot n} \cdot |z^n|^2 - 2 \cdot \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2 | k<l\}} \left\| a_i \cdot \bar{a_j} \right\| \cdot \|z\|^i \cdot \|z\|^j$$

$$\underset{0\leqslant |z^n|^2 \text{ and } \in\mathbb{R} 13.15}{=} |a_n|^{2\cdot n} \cdot \||z^n|^2\| - 2 \cdot \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2 | k<l\}} \left\| a_i \cdot \bar{a_j} \right\| \cdot \|z\|^{i+j}$$

$$= |a_n|^{2\cdot n} \cdot \|z^n \cdot \overline{z^n}\| - 2 \cdot \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2 | k<l\}} \left\| a_i \cdot \bar{a_j} \right\| \cdot \|z\|^{i+j}$$

$$\geqslant_{13.15\ (3)} |a_n|^{2\cdot n} \frac{\|z^n\| \cdot \|\overline{z^n}\|}{2} - 2 \cdot \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2 | k<l\}} \left\| a_i \cdot \bar{a_j} \right\| \cdot \|z\|^{i+j}$$

$$\underset{13.15 (2)}{=} |a_n|^{2\cdot n} \frac{\|z^n\|^2}{2} - 2 \cdot \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2 | k<l\}} \left\| a_i \cdot \bar{a_j} \right\| \cdot \|z\|^{i+j}$$

$$\geqslant_{13.15\ (7)} |a_n|^{2\cdot n} \frac{1}{2} \cdot \left(\frac{\|z\|^n}{2^n} \right)^2 - 2 \cdot \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2 | k<l\}} \left\| a_i \cdot \bar{a_j} \right\| \cdot \|z\|^{i+j}$$

$$= |a_i|^{2\cdot n} \frac{\|z\|^{2\cdot n}}{2^{2\cdot n+1}} - 2 \cdot \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2 | k<l\}} \left\| a_i \cdot \bar{a_j} \right\| \cdot \|z\|^{i+j}$$

proving that

$$(p \cdot \bar{p})(z) \geqslant |a_n|^{2\cdot n} \frac{\|z\|^{2\cdot n}}{2^{2\cdot n+1}} - 2 \cdot \sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2 | k<l\}} \left\| a_i \cdot \bar{a_j} \right\| \cdot \|z\|^{i+j} \tag{13.17}$$

Further by 13.27 that $A=\{(k,l)\in\{0,\ldots,n\}^2|k<l\}=\bigsqcup_{m\in\{1,\ldots,2\cdot n-1\}}B_m$ where $B_m=\{(k,l)\in A|k+l=m\}$ hence

$$\begin{aligned}\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}}\left\|a_i\cdot\bar{a_j}\right\|\cdot\|z\|^{i+j} &= \sum_{(i,j)\in\bigsqcup_{m\in\{1,\ldots,2\cdot n-1\}}B_m}\left\|a_i\cdot\bar{a_j}\right\|\cdot\|z\|^{i+j}\\ &\underset{10.47}{=} \sum_{m\in\{1,\ldots,2\cdot n-1\}}\left(\sum_{(i,j)\in B_m}\|a_i\cdot\bar{a_j}\|\cdot\|z\|^{i+j}\right)\\ &= \sum_{m\in\{1,\ldots,2\cdot n-1\}}\left(\sum_{(i,j)\in B_m}\|a_i\cdot\bar{a_j}\|\cdot\|z\|^{m}\right)\\ &= \sum_{m=1}^{2\cdot n-1}\left(\sum_{(i,j)\in B_m}\|a_i\cdot\bar{a_j}\|\right)\|z\|^m\\ &= \sum_{m=1}^{2\cdot n-1}b_m\cdot\|z\|^m\end{aligned}$$

where $b_m=\sum_{(i,j)\in B_m}\|a_i\cdot\bar{a_j}\|$. Using the above and 13.17 gives then that $(p\cdot\bar{p})(z)\geqslant|a_n|^{2\cdot n}\frac{\|z\|^{2\cdot n}}{2^{2\cdot n+1}}-2\cdot\sum_{m=1}^{n-1}b_m\|z\|^n$ so if we define $\{c_i\}_{i\in\{0,\ldots,n\}}$ by $c_i=\begin{cases}-2\cdot b_m\text{ if } i\in\{1,\ldots,2\cdot n-1\}\\ \frac{|a_n|^{2\cdot n}}{2^{2\cdot n+1}}\text{ if } i=2\cdot n\end{cases}$ then

$$c_n>0\wedge(p\cdot\bar{p})(z)\geqslant\sum_{m\in\{1,\ldots,2\cdot n\}}c_i\cdot\|z\|^i \tag{13.18}$$

As $c_n>0$ we have by 13.13 that $\lim_{x\to\infty}\left(\sum_{i=1}^{2\cdot n}c_i\cdot x^i\right)=\infty$ hence there exists a $R'\in\mathbb{R}$ such that if $x\geqslant R'$ then $\sum_{i=1}^{2\cdot n}c_i\cdot x^i\geqslant(p\cdot\bar{p})(0)+1$ so if $x\geqslant\max(R',1)>0\Rightarrow x\geqslant R'$ we have $\sum_{i=1}^{2\cdot n}c_i\cdot x^i>(p\cdot\bar{p})(0)$. Hence by taking $R=\max(R',1)$ we have proved that

$$\exists R>0\text{ so that if }\|z\|\geqslant R\text{ then }(p\cdot\bar{p})(z)>(p\cdot\bar{p})(0) \tag{13.19}$$

Now as $\overline{B_{\|\|}}(0,R)=\{x\in\mathbb{C}|\|x\|\leqslant\mathbb{R}\}$ is a closed and bounded set in $\langle\mathbb{C},\|\|\rangle$ (see 12.56 and 12.66) we have by 12.295 that $\overline{B_{\|\|}}(0,R)$ is compact. By the extreme value theorem (see 12.247) and the fact that $p\cdot\bar{p}$ is continuous (see 13.16 (5) we have that there exists a $z_0\in\overline{B_{\|\|}}(0,R)$ such that $(p\cdot\bar{p})(z)\geqslant(p\cdot\bar{p})(z_0)$ for every $z\in\overline{B_{\|\|}}(0,R)$. As $0\in\overline{B_{\|\|}}(0,R)$ we have also $(p\cdot\bar{p})(0)\geqslant(p\cdot\bar{p})(z_0)$. Hence if $z\in\mathbb{C}$ then we have either $z\leqslant\mathbb{R}$ and thus by the above we have $(p\cdot\bar{p})(z)\geqslant(p\cdot\bar{p})(z_0)$ or $R<z$ and then by 13.19 we have $(p\cdot\bar{p})(z)\geqslant(p\cdot\bar{p})(0)\geqslant(p\cdot\bar{p})(z_0)$, so we reach the conclusion that

$$\forall z\in\mathbb{C}\text{ we have }(p\cdot\bar{p})(z)\geqslant(p\cdot\bar{p})(z_0) \tag{13.20}$$

If we define now

$$r:\mathbb{C}\rightarrow\mathbb{C} \text{ by } r(z)=p(z+z_0) \tag{13.21}$$

then $r(0)=p(0+z_0)=p(z_0)$ and $(r\cdot\bar{r})(z)=r(z)\cdot\overline{r(z)}=p(z+z_0)\cdot\overline{p(z+z_0)}\geqslant(p(z_0)\cdot\overline{p(z_0)})=r(0)\cdot\overline{r(0)}=(r\cdot\bar{r})(0)$ so using 13.20 we have that

$$\forall z\in\mathbb{C} \text{ we have } r(z)\cdot\overline{r(z)}-r(0)\cdot\overline{r(0)}\geqslant 0 \tag{13.22}$$

By 13.23 we have that r is a non constant polynomal of order n hence using 13.24 we have

$$\exists k\in\{1,\ldots,n\} \text{ and a polynomal } q \text{ of order } n-k \text{ such that } q(0)\neq 0\wedge r(z)=z^k\cdot q(z)+r(0) \tag{13.23}$$

Then we have that

$$\begin{aligned}
r(z)\cdot\overline{r(z)}-r(0)\cdot\overline{r(0)} &= (z^k\cdot q(z)+r(0))\cdot\overline{z^k\cdot q(z)+r(0)}-r(0)\cdot\overline{r(0)}\\
&= (z^k\cdot q(z)+r(0))\cdot(\overline{z^k\cdot q(z)}+\overline{r(0)})-r(0)\cdot\overline{r(0)}\\
&= z^k\cdot q(z)\cdot\overline{z^k\cdot q(z)}+z^k\cdot q(z)\cdot\overline{r(0)}+r(0)\cdot\overline{z^k\cdot q(z)}+r(0)\cdot\\
&\quad\ \overline{r(0)}-r(0)\cdot\overline{r(0)}\\
&= z^k\cdot q(z)\cdot\overline{z^k\cdot q(z)}+z^k\cdot q(z)\cdot\overline{r(0)}+\overline{z^k\cdot q(z)\cdot\overline{r(0)}}\\
&= z^k\cdot q(z)\cdot\overline{z^k\cdot q(z)}+2\cdot\mathrm{Re}(z^k\cdot q(z)\cdot\overline{r(0)})
\end{aligned}$$

Using 13.22 we have then that

$$z^k\cdot q(z)\cdot\overline{z^k\cdot q(z)}+2\cdot\mathrm{Re}(z^k\cdot q(z)\cdot\overline{r(0)})\geqslant 0 \tag{13.24}$$

If now $x\in\mathbb{C}$ and $\delta\in\mathbb{R}$ with $0<\delta$ then substituing $\delta\cdot x$ for z in the above gives

$$\begin{aligned}
0 &\leqslant (\delta\cdot x)^k\cdot q(\delta\cdot x)\cdot\overline{(\delta\cdot x)^k\cdot q(\delta\cdot x)}+2\cdot\mathrm{Re}((\delta\cdot x)^k\cdot q(\delta\cdot x)\cdot\overline{r(0)})\\
&= \delta^k\cdot x^k\cdot q(\delta\cdot x)\cdot\delta^k\cdot\overline{x^k\cdot q(\delta\cdot x)}+\delta^k\cdot 2\cdot\mathrm{Re}(x^k\cdot q(\delta\cdot x)\cdot\overline{r(0)})\\
&= \delta^k\cdot[x^k\cdot q(\delta\cdot x)\cdot\delta^k\cdot\overline{x^k\cdot q(\delta\cdot x)}+2\cdot\mathrm{Re}(x^k\cdot q(\delta\cdot x)\cdot\overline{r(0)})]\\
&= \delta^k\cdot[\delta^k\cdot|x^k\cdot q(\delta\cdot x)|^2+2\cdot\mathrm{Re}(x^k\cdot q(\delta\cdot x)\cdot\overline{r(0)})]
\end{aligned}$$

hence dividing by $0<\delta^k$ we have then

$$\forall r>0,\forall x\in\mathbb{C} \text{ that } 0\leqslant\delta^k\cdot|x^k\cdot q(\delta\cdot x)|^2+2\cdot\mathrm{Re}(x^k\cdot q(\delta\cdot x)\cdot\overline{r(0)}) \tag{13.25}$$

If we consider the function $f_1:\mathbb{R}\rightarrow\mathbb{R}$ defined by $f_1(\delta)=\mathrm{Re}(x^k\cdot q(\delta\cdot x)\cdot\overline{r(0)})$ then using the definitions in 13.16 we have that $f_1(\delta)=\mathrm{Re}(x^k\cdot q(\delta\cdot x)\cdot\overline{r(0)})=(\mathrm{Re}\circ(\cdot(x^k\cdot\overline{r(0)}))\circ q_{\cdot x})(\delta)$ hence f_1 is continuous as the product of continuous functions (see 13.16). Further the function $f_2:\mathbb{R}\rightarrow\mathbb{R}$ defined by $f_2(\delta)=|x^k\cdot q(\delta\cdot x)|^2=(\diamondsuit\circ(\cdot(x^k))\circ q_{\cdot z})(\delta)$ is continuous as the composition of continuous functions (see 13.16). Hence $f_3:\mathbb{R}\rightarrow\mathbb{R}$ defined by $f_3(\delta)=\delta^k\cdot|x^k\cdot q(\delta\cdot x)|^2=\delta^k\cdot f_2(\delta)$ is continuous as the product of continuous function and $f_4:\mathbb{R}\rightarrow\mathbb{R}$ defined by $f_4(\delta)=2\cdot\mathrm{Re}(x^k\cdot q(\delta\cdot x)\cdot\overline{r(0)})$ is continuous as the product of a constant (thus continuous function) and a continuous function. So finally $f: \mathbb{R} \to \mathbb{R}$ defined by $f(\delta) = \delta^k \cdot |x^k \cdot q(\delta \cdot x)|^2 + 2 \cdot \mathrm{Re}(x^k \cdot q(\delta \cdot x) \cdot \overline{r(0)}) = f_3(\delta) + f_4(\delta)$ is continuous as the sum of continuous functions. To summerize

$$\forall x \in \mathbb{C} \vDash f: \mathbb{R} \to \mathbb{R} \text{ defined by } f(\delta) = \delta^k \cdot |x^k \cdot q(\delta \cdot x)|^2 + 2 \cdot \mathrm{Re}(x^k \cdot q(\delta \cdot x) \cdot \overline{r(0)}) \text{ is continuous} \tag{13.26}$$

Using 13.25 we have $\forall \delta > 0$ that $0 \leqslant f(\delta)$. Asume now that $f(0) < 0$ then if we take $\varepsilon = -f(0)$ there exists by continuity of f a $\zeta > 0$ such that for $0 < \frac{\zeta}{2} < \zeta$ we have $\left|f\left(\frac{\zeta}{2}\right) - f(0)\right| < -f(0)$ and as $f(0) < 0 \leqslant f\left(\frac{\zeta}{2}\right)$ we have $f\left(\frac{\zeta}{2}\right) - f(0) = \left|f\left(\frac{\zeta}{2}\right) - f(0)\right| < -f(0)$ proving that $0 \leqslant f\left(\frac{\zeta}{2}\right) < 0$ a contradiction. So we must have that $0 \leqslant f(0)$ giving using the definition of f that

$$\forall x \in \mathbb{C} \text{ we have } 0 \leqslant \mathrm{Re}(x^k \cdot q(0) \cdot \overline{r(0)}) = \mathrm{Re}(x^k \cdot (a + i \cdot b)) \text{ where } q(0) \cdot \overline{r(0)} = a + i \cdot b \tag{13.27}$$

Then for $k \in \{1, \ldots, n\}$ we have the following cases

***k* is odd.** then if $x = 1$ we have $0 \leqslant \mathrm{Re}(1^k \cdot (a + i \cdot b)) = \mathrm{Re}(a + i \cdot b) = a$, if $x = -1$ we have $0 \leqslant \mathrm{Re}((-1)^k \cdot (a + i \cdot b)) = \mathrm{Re}(-a - i \cdot b) = -a$ from which it follows that

$$a = 0 \tag{13.28}$$

As k is odd then we have that $k = 2 \cdot m + 1$ and for m we have again two cases

m = even. then $m = 2 \cdot l \Rightarrow k = 4 \cdot l + 1$ hence if $x = i$ we have by 9.22 that $x^k = i^k = i$ and thus $0 \leqslant \mathrm{Re}(i \cdot a - b) = -b$ and if $x = -i$ then $x^k = (-i)^k = (-1)^k \cdot i^k = -i$ and thus $0 \leqslant \mathrm{Re}(-i \cdot a + b) = b$ proving that $b = 0$

m = odd. then $m = 2 \cdot l + 1 \Rightarrow k = 4 \cdot l + 3$ hence if $x = i$ we have by 9.22 that $x^k = i^k = -i$ and thus $0 \leqslant \mathrm{Re}(-i \cdot a + b) = b$ and if $x = -i$ then $x^k = (-i)^k = (-1)^k \cdot i^k = (-1) \cdot (-i) = i$ and thus $0 \leqslant \mathrm{Re}(i \cdot a - b) = -b$ proving that $b = 0$

So in all cases we have that $b = 0$ which together with 13.28 we have that $q(0) \cdot \overline{r(0)} = 0$ we have as $q(0) \neq 0$ (see 13.23) that $\overline{r(0)} = 0$ hence we have

$$r(0) = 0$$

***k* is even.** As k is even we have $k = 2 \cdot m$ and then we have two cases for m to consider

m is odd. then $m = 2 \cdot l + 1$ and $k = 4 \cdot l + 2$, using 9.22 we have then that $i^k = -1$ so if $x = i \Rightarrow x^k = -1$ we have $0 \leqslant \mathrm{Re}(-a - i \cdot b) = -a$, if $x = 1$ then $x^k = 1$ so that $0 \leqslant \mathrm{Re}(a + i \cdot b) = a$ which proves that $a = 0$. Hence if we take $x = (1 \pm i)$ we have $0 \leqslant \mathrm{Re}(x^k \cdot (a + i \cdot b)) = \mathrm{Re}((1 \pm i)^{4 \cdot l + 2} \cdot i \cdot b) \underset{13.17}{=} \mathrm{Re}(\pm 2 \cdot (-4)^l \cdot i \cdot i \cdot b) = \mathrm{Re}(\mp 2 \cdot (-4)^l \cdot b) = \mp 2 \cdot (-4)^l \cdot b$ proving

$0 \leqslant b \vee 0 \leqslant -b$ hence $b = 0$. So we have $q(0) \cdot \overline{r(0)} = 0 \underset{q(0) \neq 0}{\Rightarrow} \overline{r(0)} = 0$ and thus

$$r(0) = 0$$

***m* is even.** then $m = 2 \cdot l$ and $k = 4 \cdot l$ if $x = 1$ we have $0 \leqslant \mathrm{Re}(1^k \cdot (a + i \cdot b)) = \mathrm{Re}(a + i \cdot b) = a$ proving that

$$0 \leqslant a \tag{13.29}$$

Use now $x = \left(1 + \frac{i}{2 \cdot l}\right)$ we have then that $x^k = \left(1 + \frac{i}{2 \cdot l}\right)^{4 \cdot l} = \left(\left(1 + \frac{i}{2 \cdot l}\right)^2\right)^k$ then using 13.28 we have that $\mathrm{Re}(x^k) < 0 < \mathrm{Im}(x^k)$. As $x^k \in \mathbb{C}$ we have a $c, d \in \mathbb{R}$ with $x^k = c + i \cdot d$ and $c = \mathrm{Re}(x^k) < 0 < \mathrm{Img}(x^k) = d$. Hence using 13.27 we have $0 \leqslant \mathrm{Re}((c + i \cdot d) \cdot (a + i \cdot b)) = \mathrm{Re}((c \cdot a - d \cdot b) + i \cdot (c \cdot b + d \cdot a)) = c \cdot a - d \cdot b$ proving

$$0 \leqslant a \cdot c - b \cdot d \tag{13.30}$$

If we take now $y = \bar{x}$ so that $y^k = \bar{x}^k \underset{9.9}{=} \overline{x^k} = c - i \cdot d$ by 13.27 we have then that $0 \leqslant \mathrm{Re}(y^k \cdot (a + i \cdot b)) = \mathrm{Re}((c - i \cdot d) \cdot (a + i \cdot b)) = \mathrm{Re}((c \cdot a + d \cdot b) + i \cdot (c \cdot b - d \cdot a)) = c \cdot a + d \cdot b$ proving

$$0 \leqslant a \cdot c + b \cdot d \tag{13.31}$$

Summing 13.30 and 13.31 gives that $0 \leqslant a \cdot c - b \cdot d + a \cdot c + b \cdot d = 2 \cdot a \cdot c \Rightarrow 0 \leqslant a \cdot c$ and as $c < 0$ we have $a \leqslant 0$ which by 13.29 proves that $a = 0$. Substituting this in 13.30 and 14.135 proves that $0 \leqslant -b \cdot d \Rightarrow b \cdot d \leqslant 0$ and $0 \leqslant b \cdot d$ hence $b \cdot d = 0$ and as $0 < d$ we have $b = 0$. So we concude that $q(0) \cdot \overline{r(0)} = 0 \underset{q(0) \neq 0}{\Rightarrow} \overline{r(0)} = 0$ and thus

$$r(0) = 0$$

hence for all cases of k we have that $r(0) = 0$. As by definition (see 13.21) $0 = r(0) = p(0 + z_0) = p(z_0)$ we have proved finally that $p(z_0) = 0$ and we have find a zero of the polyomial p proving the fundamental theorem of algebra. □

Theorem 13.30. *Let p be a polynomial with real coefficients and $z \in \mathbb{C}$ is such that $p(z) = 0$ then $p(\bar{z}) = 0$*

Proof. If p is defined by the coefficients $\{a_i\}_{i \in \{0, \ldots, n\}} \subseteq \mathbb{R}$ then we have as $p(z) = 0$ that

$$\begin{aligned} 0 &= \bar{0} \\ &= \overline{p(z)} \\ &= \overline{\sum_{i=0}^{n} a_i \cdot z^i} \end{aligned}$$

$$
\begin{aligned}
&\underset{13.25}{=} \sum_{i=0}^{n} \overline{a_i \cdot z^i} \\
&= \sum_{i=0}^{n} \bar{a_i} \cdot \bar{z^i} \\
&\underset{9.9}{=} \sum_{i=0}^{n} \bar{a_i} \cdot \bar{z}^i \\
&= \sum_{i=0}^{n} a_i \cdot \bar{z}^i \\
&= p(\bar{z})
\end{aligned}
$$

proving the theorem. □

Lemma 13.31. *Let p be a polynomial then $p(x_0)=0$ if and only if there exists a polynomial q such that $\forall x \in \mathbb{C}$ we have $p(x)=(x-x_0)\cdot q(x)$ and* $\mathrm{ord}(q)=\mathrm{ord}(p)-1$

Proof.

⇒. Suppose $p(x_0)=0$ then we have to consider the following cases

$\mathrm{ord}(p)=0$. then $p(x)=\sum_{i=0}^{0} a_0 \cdot x^i = a_0$ hence as $a_0=p(x_0)=0$ we have $a_0=0$ so if we define q by $q=C_0$ we have $p(x)=a_0=0=(x-x_0)\cdot p(x)=(x-x_0)\cdot 0$

$\mathrm{ord}(p)=1$. then $p(x)=\sum_{i=0}^{1} a_i \cdot x^i = a_1 \cdot x + a_0$ with $a_1 \neq 0$, hence as $0=p(x_0)=a_1 \cdot x_0 + a_0$ we have $a_0=-a_1 \cdot x_0$ hence $p(x)=a_1 \cdot x - a_1 \cdot x_0=(x-x_0)\cdot a_1=(x-x_0)\cdot q(x)$ where $q(x)=a_1 \neq 0$ so that $\mathrm{ord}(q)=0$

$\mathrm{ord}(p)>1$. then if $n=\mathrm{ord}(p)$ we have $p(x)=\sum_{i=0}^{n} a_i \cdot x^i$ and $a_n \neq 0$ so that $0=p(x_0)=\sum_{i=0}^{n} a_i \cdot x^i=\sum_{i=1}^{n} a_i \cdot (x_0)^i + a_0 \Rightarrow a_0=-\sum_{i=1}^{n} a_i \cdot (x_0)^i$ so that

$$
\begin{aligned}
p(x) &= \sum_{i=0}^{n} a_i \cdot x^i \\
&= \sum_{i=1}^{n} a_i \cdot x^i + a_0 \\
&= \sum_{i=1}^{n} a_i \cdot x^i + \sum_{i=1}^{n} a_i \cdot (x_0)^i \\
&= \sum_{i=1}^{n} a_i \cdot (x^i - (x_0)^i) \\
&= (x-x_0) \cdot \sum_{i=1}^{n} a_i \cdot (x^i - (x_0)^{i-1})
\end{aligned}
$$

$$\begin{aligned}
&= (x-x_0)\cdot\sum_{i=0}^{n-1} a_{i+1}\cdot(x^i-(x_0)^i)\\
&= (x-x_0)\cdot\left(\sum_{i=1}^{n-1} a_{i+1}\cdot x^i+\left(a_1-\sum_{i=0}^{n-1} a_{i+1}\cdot(x_0)^i\right)\right)\\
&= (x-x_0)\cdot q(x)
\end{aligned}$$

where $q=\sum_{i=0}^{n-1} b_i\cdot x^i$ and $b_i=\begin{cases}(a_1-\sum_{j=0}^{n-1} a_{j+1}\cdot(x_0)^j) \text{ if } i=0\\ a_{i+1} \text{ if } i\in\{1,\ldots,n-1\}\end{cases}$ as $b_{n-1}=a_n\neq 0$ we have also $\operatorname{ord}(q)=n-1=\operatorname{ord}(p)-1$

$\Leftarrow$. If $p(x)=(x-x_0)\cdot q(x)$ then $p(x_0)=(x_0-x_0)\cdot q(x_0)=0\cdot q(x_0)=0$ □

Theorem 13.32. *Let p be a non constant polynomial then there exists a $0\neq c\in\mathbb{C}$, $\{x_i\}_{i\in\{1,\ldots,\operatorname{ord}(p)\}}$ such that $\forall x\in\mathbb{C}$ we have $p(x)=c\cdot\prod_{j=1}^{\operatorname{ord}(p)}(x-x_i)$. Furthermore we have that $\{z\in\mathbb{C}|p(z)=0\}=\{x_i|i\in\{1,\ldots,\operatorname{ord}(p)\}\}$*

Proof. We prove this by induction on the order of the polynomial let $\boldsymbol{S}=\{n\in\mathbb{N}|$if p is a non constant polynomial with $\operatorname{ord}(p)=n$ then $\exists c,\{x_i\}_{i\in\{1,\ldots,\operatorname{ord}(p)\}}$ such that $p(x)=c\cdot\prod_{j=1}^{\operatorname{ord}(p)}(x-x_j)\}$, then we have

- **$n=1$.** If p is a non constant polynomial with $\operatorname{ord}(p)\leqslant 1$ then $p(x)=a_1\cdot x+a_0$ where $a_1\neq 0$ hence $a_1\cdot\left(x-\left(-\frac{a_0}{a_1}\right)\right)=a_1\cdot x+a_0$ proving that $1\in\mathcal{S}$

- **$n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$.** If p is a polynomial of order $n+1$ then using the fundamental theorem of algebra (see 13.29) there exists a $x_{n+1}\in\mathbb{C}$ such that $p(x_1)=0$, applying then the previous lemma (see 13.31) there exists a q with $\operatorname{ord}(q)=n$ such that $p(x)=(x-x_{n+1})\cdot q(x)$. As $n\in\mathcal{S}$ we have that $\exists c\in\mathbb{C}, c\neq 0$, $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{C}$ such that $q(x)=c\cdot\prod_{i=1}^{n}(x-x_i)$ hence $p(x)=c\cdot(x-x_{n+1})\cdot\prod_{i=1}^{n}(x-x_i)=c\cdot\prod_{i=1}^{n+1}(x-x_i)$ proving that $n+1\in\mathcal{S}$.

Further we have if $x\in\{x_i|i\in\{1,\ldots,\operatorname{ord}(p)\}\}$ there exists a $i\in\{1,\ldots,\operatorname{ord}(P)\}$ such that $x=x_i$ hence $p(x)=p(x_i)=c\cdot\prod_{j\in\{1,\ldots,\operatorname{ord}(p)\}}(x_i-x_j)=c\cdot(x_i-x_i)\cdot\prod_{j\in\{1,\ldots,\operatorname{ord}(p)\}\setminus\{i\}}(x_i-x_j)=0$ proving that $x\in\{z|p(z)=0\}$. If $x\in\{z|p(z)=0\}$ then $p(x)=0$ hence $0=c\cdot\prod_{i=1}^{\operatorname{ord}(p)}(x-x_i)\underset{c\neq 0}{\Rightarrow}\prod_{i=1}^{\operatorname{ord}(p)}(x-x_i)=0$, if now for $\forall i\in\{1,\ldots,\operatorname{ord}(p)\}$ we have that $x\neq x_i$ we have by 10.56 that $\prod_{i=1}^{\operatorname{ord}(p)}(x-x_i)\neq 0$ a contradiction so there exists a $i\in\{1,\ldots,\operatorname{ord}(p)\}$ such that $x=x_i$ and thus $x\in\{x_i|i\in\{1,\ldots,\operatorname{ord}(p)\}\}$. □

Theorem 13.33. *Let $p\in\mathcal{P}$ be a non constant polynomial then there exists a m, $M\in\mathbb{N}_0$, $c\in\mathbb{R}$ $\{\lambda_i\}_{i\in\{1,\ldots,m\}}\subseteq\mathbb{R}$, $\{b_i\}_{i\in\{1,\ldots,M\}}\subseteq\mathbb{R}$ and $\{c_i\}_{i\in\{1,\ldots,M\}}\subseteq\mathbb{R}$ such that $\forall i\in\{1,\ldots,M\}$ we have $b_i^2<4\cdot c_i$ and $\forall z\in\mathbb{C}$ we have $p(z)=c\cdot(\prod_{i=1}^{m}(z-\lambda_i))\cdot(\prod_{i=1}^{M}(x^2+b_i\cdot x+c_i))$.*

Proof. We prove this by induction so let $\mathcal{S}=\{n\in\mathbb{N}_0 \mid \text{let } p\in\mathcal{P} \text{ with } \operatorname{ord}(p)\leqslant n \text{ and real coefficients then } p(z)=(\prod_{i=1}^{m}(z-\lambda_i))\cdot(\prod_{i=1}^{M}(x^2+b_i\cdot x+c_i))\,\forall z\in\mathbb{C} \text{ where } \{\lambda_i\}_{i\in\{1,\ldots,m\}}\subseteq\mathbb{R},\ \{b_i\}_{i\in\{1,\ldots,M\}}\subseteq\mathbb{R} \text{ and } \{c_i\}_{i\in\{1,\ldots,M\}}\subseteq\mathbb{R} \text{ such that } \forall i\in\{1,\ldots,M\} \text{ we have } b_i^2<4\cdot c_i\}$ then

$1\in\boldsymbol{\mathcal{S}}$. If $\operatorname{ord}(p)=1$ then $\forall z\in\mathbb{C}$ we have $p(z)=a'\cdot z+c'=a'\cdot\left(z-\left(-\frac{c'}{a'}\right)\right)$ where $a',c'\in\mathbb{R}$ and $a'\neq 0$ hence if we take $c=a'$, $\{\lambda_i\}_{i\in\{1,\ldots,1\}}$ defined by $\lambda_1=-\frac{c'}{a'}$ and $\{b_i\}_{i\in\{1,\ldots,0\}}=\{c_i\}_{i\in\{1,\ldots,0\}}=\emptyset$ so that $\forall i\in\{1,\ldots,0\}=\emptyset$ $b_i^2<4\cdot c_i$ is satisfied vacuously. This proves that $1\in\boldsymbol{S}$

$n\in\boldsymbol{\mathcal{S}}\Rightarrow n+1\in\boldsymbol{\mathcal{S}}$. So let $p\in\mathcal{P}$ with real coefficients then if $\operatorname{ord}(p)<n+1$ then $\operatorname{ord}(p)\leqslant n$ and as $n\in\mathcal{S}$ we have that $p(z)=(\prod_{i=1}^{m}(z-\lambda_i))\cdot(\prod_{i=1}^{M}(x^2+b_i\cdot x+c_i))$, hence we must only consider the case with $\operatorname{ord}(p)=n+1$. As p is non constant we have using the previous theorem that there exists a $\{\alpha_i\}_{i\in\{1,\ldots,\operatorname{ord}(p)\}}$ such that $\forall z\in\mathbb{C}$ we have $p(z)=c\cdot\prod_{i\in\{1,\ldots,\operatorname{ord}(p)\}}(z-\alpha_i)$. If now $\forall i\in\{1,\ldots,\operatorname{ord}(p)=n+1\}$ we have $\alpha_i\in\mathbb{R}$ then $n+1\in\mathcal{S}$ by taking $m=\operatorname{ord}(p)=n$ and $M=0$. So let assume that $\exists i\in\{1,\ldots,\operatorname{ord}(p)=n+1\}$ such that $\alpha_i\in\mathbb{C}\setminus\mathbb{R}$ then $p(\alpha_i)=0$. Using 13.30 we have that $p(\bar{\alpha_i})=0$ and thus by 13.32 there exists a $j\in\{1,\ldots,\operatorname{ord}(p)=n+1\}$ such that $\bar{\alpha_i}=\alpha_j$ hence

$$\begin{aligned}
p(z) &= c\cdot\prod_{k=1}^{n+1}(z-\alpha_k)\\
&= c\cdot\prod_{k\in\{1,\ldots,n+1\}}(z-\alpha_k)\\
&= c\cdot\left(\prod_{k\in\{1,\ldots,n+1\}\setminus\{i,j\}}(z-\alpha_k)\right)\cdot\prod_{k\in\{i,j\}}(z-\alpha_k)\\
&= c\cdot\left(\prod_{k\in\{1,\ldots,n+1\}\setminus\{i,j\}}(z-\alpha_k)\right)\cdot(z-\alpha_i)\cdot(z-\alpha_j)\\
&= c\cdot\left(\prod_{k\in\{1,\ldots,n+1\}\setminus\{i,j\}}(z-\alpha_k)\right)\cdot(z-\alpha_i)\cdot(z-\alpha_j)\\
&= c\cdot\left(\prod_{k\in\{1,\ldots,n+1\}\setminus\{i,j\}}(z-\alpha_k)\right)\cdot(z-\alpha_i)\cdot(z-\bar{\alpha_i})\\
&= c\cdot\left(\prod_{k\in\{1,\ldots,n+1\}\setminus\{i,j\}}(z-\alpha_k)\right)\cdot(z^2-(\alpha_i+\bar{\alpha_i})\cdot z+\alpha_i\cdot\bar{\alpha_i})\\
&= c\cdot\left(\prod_{k\in\{1,\ldots,n+1\}\setminus\{i,j\}}(z-\alpha_k)\right)\cdot(z^2-2\cdot\operatorname{Re}(\alpha_i)\cdot z+|\alpha_i|^2)
\end{aligned}$$

Hence if we define $q(z)=c\cdot\left(\prod_{k\in\{1,\ldots,\operatorname{ord}(p)\}\setminus\{i,j\}}(x-\alpha_k)\right)$ (which is a polynomial with $\operatorname{ord}(q)=n+1-2=n-1<n$ (see 13.10)), so

$$p(z)=q(z)\cdot(z^2+b\cdot z+c) \text{ where } q\in\mathcal{P},\ \operatorname{ord}(q)<n+1 \text{ and } b=-2\cdot\operatorname{Re}(\alpha_i)\in\mathbb{R},\ c=|\alpha_i|^2\in\mathbb{R} \tag{13.32}$$

Now as $\alpha_i \in \mathbb{C} \backslash \mathbb{R}$ we have that $0 < \mathrm{Img}(\alpha_i)$ giving $(\mathrm{Re}(\alpha_i))^2 < (\mathrm{Re}(\alpha_i))^2 + (\mathrm{Img}(\alpha))^2 = |\alpha_i|^2$ hence $b^2 = 4 \cdot \mathrm{Re}(\alpha_i)^2 < 4 \cdot |\alpha_i|^2 = 4 \cdot c$ proving that

$$b^2 < 4 \cdot c \tag{13.33}$$

As q is a polynomial of order $n-1$ we have that $q(z) = \sum_{k=0}^{n-1} d_k \cdot z^k$ Now given $x \in \mathbb{R}$ we have $x^2 + b \cdot x + c = \left(x + \frac{b}{2}\right)^2 + \left(c - \frac{b^2}{4}\right) > 0$ (see above) hence using 13.32 we have that $\forall x \in \mathbb{R}$ that $q(x) = \frac{p(x)}{x^2 + b \cdot x + c} \in \mathbb{R}$ (as p has real coefficients). So $\forall x \in \mathbb{R}$ we have $0 = \mathrm{Img}(q(x)) = \mathrm{Img}(\sum_{k=0}^{n-1} d_k \cdot z^k) = \sum_{k=0}^{n-1} \mathrm{Img}(d_k \cdot x^k) \underset{x \in \mathbb{R}}{=} \sum_{k=0}^{n-1} \mathrm{Img}(d_k) \cdot z^k$ hence using 13.2 we have that $\forall k \in \{0, \ldots, n-1\}$ $\mathrm{Img}(d_k) = 0$ proving that q has real coefficients. As $n \in \mathcal{S}$ there exists then $m, M \in \mathbb{N}_0$, $\{\lambda_k\}_{k \in \{1, \ldots, m\}} \subseteq \mathbb{R}$, $\{b'_k\}_{k \in \{1, \ldots, M\}} \subseteq \mathbb{R}$ and $\{c'_k\}_{k \in \{1, \ldots, M\}} \subseteq \mathbb{R}$ such that $\forall k \in \{1, \ldots, M\}$ ${b'_k}^2 \leqslant 4 \cdot c'_k$ and $\forall z \in \mathbb{C}$ we have $q(z) = (\prod_{k=1}^m (z - \lambda_k)) \cdot (\prod_{k=1}^M (z^2 + b'_k \cdot z + c'_k))$. Hence if we define $\{b_k\}_{k \in \{1, \ldots, M+1\}}, \{c_k\}_{k \in \{1, \ldots, M+1\}} \subseteq \mathbb{R}$ by $b_k = \begin{cases} b'_k \text{ if } k \in \{1, \ldots, M\} \\ b_{M+1} = b \end{cases}$ and $c_k = \begin{cases} c'_k \text{ if } k \in \{1, \ldots, M\} \\ c_{M+1} = c \end{cases}$ then $\forall k \in \{1, \ldots, M+1\}$ we have $b_k^2 < 4 \cdot c_k$ and $\forall z \in \mathbb{C}$ that $p(z) = (\prod_{k=1}^m (z - \lambda_k)) \cdot (\prod_{k=1}^{M+} (z^2 + b_k \cdot z + c_k))$, finally proving that $n + 1 \in \mathcal{S}$ $\square$

13.2 Spectral theorem

13.2.1 Polynomials and operators

We use now the fundamental theorem of algebra to prove the spectral theorem of linear algebra. First we define the concept of operators

Definition 13.34. (Linear operator) *Let $\langle X, +, \cdot \rangle$ be a vector space then $L: X \rightarrow X$ is a **linear operator on X** if $L \in \mathrm{Hom}(X, X)$ (see 10.184). The set of all linear operators oon X is called* $\mathrm{Hom}(X)$.

Theorem 13.35. *Let $\langle X, +, \cdot \rangle$ be a non trivial finite dimensional vector space, then for $L \in \mathrm{Hom}(X)$ we have the following equivalences*

1. *L is a isomorphism*
2. *L is injective*
3. *L is surjective*

Proof.

1 ⇒ 2. *This is trivial as a isomorphism is bijective and thus injective*

1 ⇒ 3. *This is trivial as a isomorphism is bijective and thus surjective*

2 ⇒ 1. *As L is injective we have by 10.220 that* $\operatorname{rank}(L) = \dim(X)$ *and by 10.220 again (using the fact that $L: X \Rightarrow X$) that L is a isomorphim.*

3 ⇒ 1. *As L is surjective we have by 10.217 that* $\operatorname{rank}(L) = \dim(X)$ *and by 10.220 (using the fact that $L: X \Rightarrow X$) that L is a isomorphim.* □

Definition 13.36. *Let $\langle X, +, \cdot\rangle$ a vactor space and $\{L_i: X \to X\}_{i\in\{1,\ldots,.n\}}$ a familly of operators between X then $\dot{\prod}_{i=1}^n L_i: X \to X$ is defined recursively by* $(\dot{\prod}_{i=1}^n L_i)(x) = \begin{cases} L_1(x) \text{ if } n=1 \\ L_n((\dot{\prod}_{i=1}^{n-1} L_i)(x)) \text{ if } n>1 \end{cases}$

Definition 13.37. *Let X be a vector space and $L \in \operatorname{Hom}(X)$ then given $n \in \mathbb{N}_0$ we define $L^n: X \to X$ recusively by* $L^n = \begin{cases} 1_X \text{ if } n=0 \\ L \circ L^{n-1} \text{ if } n>0 \end{cases}$

Note 13.38. As a application of the above definition we have we have that $L^0 = 1_X$ and $L^1 = L$

The power of operators behaves as a normal powerfunction.

Lemma 13.39. *Let $\langle X,+,\cdot\rangle$ be a vector space and $L \in \operatorname{Hom}(X)$ then given n, $m \in \mathbb{N}_0$ we have $L^n \circ L^m = L^{n+m}$*

Proof. We prove this by induction on n so let $\mathcal{S} = \{n \in \mathbb{N}_0 | L^n \circ L^m = L^{n+m}\}$ then we have

$0 \in \mathcal{S}$. this follows from $L^0 \circ L^m = 1_X \circ L^m = L^m = L^{0+m}$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$. then $L^{n+1} \circ L^m = (L \circ L^n) \circ L^m \underset{\text{composition is associative}}{=} L \circ (L^n \circ L^m) \underset{n \in \mathcal{S}}{=} L \circ L^{n+m} = L^{(n+1)+m}$ proving that $n+1 \in \boldsymbol{S}$ □

The power of a linear operator is again a linear operator

Lemma 13.40. *Let $\langle X, +, \cdot\rangle$ be a vector space over a field $\mathcal{F}$, $L \in \operatorname{Hom}(X)$ and $n \in \mathbb{N}_0$ then $L^n \in \operatorname{Hom}(X)$ (L^n is a linear operator)*

Proof. Again this easely proved by induction and define $\mathcal{S} = \{n \in \mathbb{N}_0 | L^n \text{ is linear}\}$ then we have

$0 \in \mathcal{S}$. then $L^0 = 1_X$ which is linear so that $0 \in \mathcal{S}$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$. As $L^{n+1} = L \circ L^n$ which is linear as L is linear, L^n is linear and the composition of linear maps is linear (see 10.182) proving that $n+1 \in \mathcal{S}$ □

Using the definition of the power of a operator we can define a polynomial of linear operator

Definition 13.41. *Let $n\in\mathbb{N}_0$, $\langle X,+,\cdot\rangle$ a complex vector space and $p:\mathbb{C}\rightarrow\mathbb{C}$ a polynomial (see 13.1) defined by $\{\alpha_i\}_{i\in\{1,\ldots,n\}}$ and $L\in\mathrm{Hom}(X)$ a linear operator then $p(L):X\rightarrow X$ is defined by $p(E)(x)=\sum_{i=0}^{n} L^i(x)$. $p(L)$ is called a polynomial operator.*

We show now that the product of polynomial functions is equal to the composition of the operator polynomials.

Lemma 13.42. *Let $\langle X,+,\cdot\rangle$ a complex vector space and $p,q:\mathbb{C}\rightarrow\mathbb{C}$ polynomials and $L\in\mathrm{Hom}(X)$ then $(p\cdot q)(L)=p(L)\circ q(L)$*

Proof. Let p be defined $p(z)=\sum_{i=0}^{n} a_i\cdot z^i$ and q defined by $q(z)=\sum_{i=0}^{m} b_i\cdot z^i$ we have by 13.8 that

$$(p\cdot q)(z)=\sum_{i=0}^{n+m}\left(\sum_{(k,l)\in\{(r,s)\in\{0,\ldots,n\}\times\{0,\ldots,m\}|r+s=i\}} a_k\cdot b_l\right)\cdot z^n \tag{13.34}$$

Next we calculate $p(L)\circ q(L)$ first we prove that given $x\in X$ we have

$$(p(L)\circ q(L))(x)=\sum_{(k,l)\in\{0,\ldots,n\}\times\{0,\ldots,m\}} a_k\cdot b_l\cdot L^{k+k}(x) \tag{13.35}$$

To prove this note that

$$\begin{aligned}
(p(L)\circ q(L))(x) &= p(L)(q(L)(x))\\
&= \sum_{i=0}^{n} a_i\cdot L^i(q(L)(x))\\
&= \sum_{i=1}^{n} a_i\cdot L^i\!\left(\sum_{j=0}^{m} b_j\cdot L^j(x)\right)\\
&\underset{L^i \text{ is linear by } 13.40}{=} \sum_{i=0}^{n} L^i\!\left(a_i\cdot\sum_{j=0}^{m} b_j\cdot L^j(x)\right)\\
&= \sum_{i=0}^{n} L^i\!\left(\sum_{j=0}^{m} a_i\cdot b_j\cdot L^j(x)\right)\\
&\underset{10.174 \text{ and } 13.40}{=} \sum_{i=0}^{n}\left(\sum_{j=0}^{m} a_i\cdot b_j\cdot L^i(L^j(x))\right)
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=0}^{n}\left(\sum_{j=0}^{m} a_i\cdot b_j\cdot (L^i\circ L^j)(x)\right)\\
&\underset{13.39}{=} \sum_{i=0}^{n}\left(\sum_{j=0}^{m} a_i\cdot b_j\cdot L^{i+j}(x)\right)\\
&= \sum_{i\in\{0,\ldots,n\}}\left(\sum_{j\in\{0,\ldots,m\}} a_i\cdot b_j\cdot L^{i+j}(x)\right)\\
&\underset{10.48}{=} \sum_{(k,l)\in\{0,\ldots,n\}\times\{0,\ldots,m\}} a_k\cdot b_l\cdot L^{k+k}(x)
\end{aligned}$$

which proves 13.35. Now using 13.7 we have that $\{0,\ldots,n\}\times\{0,\ldots,m\}$ is a disjoint union of the sets $\{(r,s)\subseteq\{0,\ldots,n\}\times\{0,\ldots,m\}|r+s=i\}$ where $i\in\{0,\ldots,n+m\}$, so using 13.35 together with 10.47 we have

$$\begin{aligned}
(p(L)\circ q(L))(x) &= \sum_{i\in\{0,\ldots,n+m\}}\left(\sum_{(k,l)\in\{(r,s)\subseteq\{0,\ldots,n\}\times\{0,\ldots,m\}|r+s=i\}} a_k\cdot b_l\cdot L^{k+k}(x)\right)\\
&= \sum_{i\in\{0,\ldots,n+m\}}\left(\sum_{(k,l)\in\{(r,s)\subseteq\{0,\ldots,n\}\times\{0,\ldots,m\}|r+s=i\}} a_k\cdot b_l\cdot L^{i}(x)\right)\\
&= \sum_{i\in\{0,\ldots,n+m\}}\left(\sum_{(k,l)\in\{(r,s)\subseteq\{0,\ldots,n\}\times\{0,\ldots,m\}|r+s=i\}} a_k\cdot b_l\right)\cdot L^{i}(x)\\
&\underset{13.34}{=} (p\cdot q)(L)(x)
\end{aligned}$$

which as $x\in X$ is choosen arbitrary proves that $(p\cdot q)(L)=p(L)\circ q(L)$ $\square$

We extend the above by recursion to the composition of n linear operators

Corollary 13.43. *Let $n\in\mathbb{N}$, $\langle X,+,\cdot\rangle$ a complex vector space and $L\subseteq\mathrm{Hom}(X)$ and $\{p_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathcal{P}$ a finite non empty family of polynomials (hence by 13.10 $\prod_{i=1}^{n}p_i$ is a polynomial) then $(\prod_{i=1}^{n}p_i)(L)=\dot{\prod}_{i=1}^{n}p_i(L)$ proving that $1\in\mathcal{S}$*

Proof. This is easely proved by induction using the previous lemma. So let $\mathcal{S}=\{n\in\mathbb{N}|\text{If }\{p_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathcal{P}\text{ then }(\prod_{i=1}^{n}p_i)(L)=\dot{\prod}p_i(L)\}$ then we have

$1\in\mathcal{S}$. then $(\prod_{i=1}^{1}p_i)(L)=p_1(L)=\dot{\prod}_{i=1}^{1}p_i(L)$

$\boldsymbol{n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}}$**.** if $\{p_i\}_{i\in\{1,\ldots,n+1\}} \subseteq \mathcal{P}$ then

$$\begin{aligned}\left(\prod_{i=1}^{n+1} p_i\right)(L) &= \left(p_{n+1}\cdot\prod_{i=1}^{n} p_i\right)(L)\\ &\underset{13.42}{=} p_{n+1}(L)\circ\left(\prod_{i=1}^{n} p_i\right)(L)\\ &\underset{n\in\mathcal{S}}{=} p_{n+1}(L)\circ\left(\prod_{i=1}^{n} p_i(L)\right)\\ &= \prod_{i=1}^{n+1} p_i(L)\end{aligned}$$

proving that $n+1\in\mathcal{S}$ □

Definition 13.44. *Let $\langle X,+,\cdot\rangle$ be a complex vector space, $L\in\mathrm{Hom}(X)$ then $\mathcal{P}[L]=\{p(L)|p\in\mathcal{P}\}$ where $\mathcal{P}$ is the set of all complex polynomials. (see 13.1)*

Using the composition as a addtive operator we can form a additive semi group

Theorem 13.45. *Let $\langle X,+,\cdot\rangle$ be a complex vector space, $L\in\mathrm{Hom}(X)$ then $\langle\mathcal{P}[L],\circ\rangle$ is a abelian semi group*

Proof. First if $p_1(L), p_2(L)\in\mathcal{P}[L]$ then $p_1(L)\circ p_2(L) \underset{13.42}{=} (p_1\cdot p_2)(L)\in\mathcal{P}[L]$ we have that $\circ\colon\mathcal{P}[L]\times\mathcal{P}[L]\to\mathcal{P}[L]$ further we have that $\circ$ as the composition of functions is associative, $1_X=\sum_{i=0}^{0} 1\cdot L^i$ is the neutral element and $p_1(L)\circ p_2(L)=(p_1\cdot p_2)(L)=(p_2\cdot p_1)(L)=p_2(L)\circ p_1(L)$ □

Combining the above with the commutativity of a sum for semi groups (see 10.19) we have the following

Theorem 13.46. *Let $\langle X,+,\cdot\rangle$ be a complex vector space, $L\in\mathrm{Hom}(X)$, $n\in\mathbb{N}$, $\{p_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathcal{P}$ and $\sigma\colon\{1,\ldots,n\}\to\{1,\ldots,n\}$ a permutation then $\dot{\prod}_{i=1}^{n} p_i(L)=\dot{\prod}_{i=1}^{n} p_{\sigma(i)}(L)$*

We are now read to prove the following theorem

Theorem 13.47. *Let $\langle X,+,\cdot\rangle$ be a complex vector space, $L\in\mathrm{Hom}(X)$ and $p\in\mathcal{P}$ a polynomial with $\mathrm{ord}(p)\in\mathbb{N}$ (a non constant polynomial) then there exists a $0\neq c\in\mathbb{C}$ and $\{\lambda_i\}_{i\in\{1,\ldots,n\}}$ such that $p(L)=c\cdot\dot{\prod}_{i=1}^{n}(L-\lambda_i\cdot L)$*

Proof. Using 13.32 there exists a $0\neq c\in\mathbb{C}$ and $\{\lambda_i\}_{i\in\{1,\ldots,n\}}$ such that $p(z)=c\cdot\dot{\prod}_{i=1}^n(z-\lambda_i)=C_c(z)\cdot\dot{\prod}_{i=1}^n(z-\lambda_i)$, now $\forall i\in\{1,\ldots,n\}$ we have $z-\lambda_i=\sum_{j=0}^1 c_{i,j}\cdot z^i$ where $c_{i,0}=-\lambda_i$ and $c_{i,1}=1$ so that $(z-\lambda_i)(L)=c_{i,1}\cdot L^1+c_{i,0}L^0=L-\lambda_i\cdot 1_X$. As c is the polynomial C_c defined by $\sum_{i=0}^0 d_i\cdot z$ where $d_i=0$ so that $C_c(L)=c\cdot 1_X$. Hence we have that

$$\begin{aligned} p(z)(L) &\underset{13.42}{=} c_c(L)\circ\left(\dot{\prod}_{i=1}^n(z-\lambda_i)\right)(L)\\ &\underset{13.43}{=} c\cdot 1_X\circ\left(\dot{\prod}_{i=1}^n(z-\lambda_i)(L)\right)\\ &= c\cdot\dot{\prod}_{i=1}^n(L-\lambda_i\cdot 1_X)\end{aligned}$$

□

Lemma 13.48. *Let $\langle X,+,\cdot\rangle$ be a vector space over a field $\mathcal{F}$, $n\in\mathbb{N}$ and $\{L_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathrm{Hom}(X)$ such that $\forall i\in\{1,\ldots,n\}$ we have that $\ker(L_i)=\{0\}$ (meaning by 10.211 that L_i is regular (=injective)) then $\ker(\dot{\prod}_{i=1}^n L_i)=\{0\}$*

Proof. We prove this by induction so let $\mathcal{S}=\{n\in\mathbb{N}|$ if $\{L_i\}_{i\in\{1,\ldots,n\}}\in\mathrm{Hom}(X)$ with $\forall i\in\{1,\ldots,n\}\vdash\ker(L_i)=\{0\}$ then $\ker(\dot{\prod}_{i=1}^n L_i)=\{0\}\}$ then we have

$1\in\mathcal{S}$. then $\dot{\prod}_{i=1}^1 L_i=L_1$ then $\ker(\dot{\prod}_{i=1}^1 L_i)=\ker(L_1)$

$n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$. then assume that $x\in\ker(\dot{\prod}_{i=1}^{n+1}L_i)$ then we have $0=(\dot{\prod}_{i=1}^{n+1}L_i)(x)=(L_{n+1}\circ\dot{\prod}_{i=1}^n L_i)(x)=L_{n+1}((\dot{\prod}_{i=1}^n L_i)(x))$ and as $\ker(L_{n+1})=0$ we have that $(\dot{\prod}_{i=1}^n L_i)(x)=0$ so that as $n\in\mathcal{S}$ we conclude that $x=0$ hence $\ker(\dot{\prod}_{i=1}^{n+1}L_i)=\{0\}$ proving that $n+1\in\mathcal{S}$ □

13.2.2 Eigenvectors and eigen values

Definition 13.49. *Let $\langle X,+,\cdot\rangle$ be a vector space over a field $\mathcal{F}$, $L\in\mathrm{Hom}(X)$ then x is a **eigenvector** of L with* **eigenvalue** *λ iff $x\neq 0$ and $L(x)=\lambda\cdot x$*

Remark 13.50. The condition that $x\neq 0$ is needed because we have $L\cdot 0=0=\lambda\cdot 0$ for every $\lambda\in\mathcal{F}$ so the concept of a special eigen value does not make sense.

In a complex finite dimensional vector space every linear operator has at least one eigen vector as is expressed in the following theorem.

Theorem 13.51. *Let $\langle X,+,\cdot\rangle$ be a non trivial complex finite dimensional non trivial vector space then every $L\in\mathrm{Hom}(X)$ has a eigen vector with a eigen value.*

Proof. Let $n=\dim(X)\in\mathbb{N}$ then as X is non trivial there exists a $x\in X$ with $x\neq 0$, consider then the family $\{L^i(x)\}_{i\in\{0,\ldots,n\}}$ of $n+1$ vectors in X which as $\dim(X)$ is n must be linear dependent. Hence there exists a $\{a_i\}_{i\in\{0,\ldots,n\}}$ not all equal to 0 such that $\sum_{i=0}^n a_i\cdot L^i(x)=0$. If we take then $m=\max(\{i\in\{0,\ldots,n\}|a_i\neq 0\})$ then we have that

$$\sum_{i=0}^m a_i\cdot L^i(x)=0\wedge a_m\neq 0 \tag{13.36}$$

If now $m=0$ then $0=\sum_{i=0}^{m} a_i \cdot L^i(x) = a_0 \cdot L^0(x) = a_0 \cdot x = 0 \underset{a_0 \neq 0}{\Rightarrow} x=0$ a contradiction hence we have $m>0$ and $p \in \mathcal{P}$ defined by $p(z) = \sum_{i=0}^{m} a_i \cdot z^i$ is a non constant polynomial. Applying then 13.47 we find that there exists a $0 \neq c \in \mathbb{C}$ and $\{\lambda_i\}_{i \in \{0,\ldots,m\}} \subseteq \mathbb{C}$ we have that

$$\sum_{i=0}^{m} a_i \cdot L^i = c \cdot \prod_{i=1}^{m} (L - \lambda_i \cdot 1_X) \tag{13.37}$$

So we have

$$\begin{aligned} 0 &\underset{13.36}{=} \sum_{i=0}^{m} a_i \cdot L^i(x) \\ &\underset{13.37}{=} c \cdot \left(\prod_{i=1}^{m} (L - \lambda_i \cdot 1_X) \right)(x) \end{aligned}$$

Hence as $v \neq 0$ we have that $c \neq 0$ we have that $\left(\prod_{i=1}^{m} (L - \lambda_i \cdot 1_X)\right)(v) = 0$ and as $v \neq 0$ we have using 13.48 that there exists a $i \in \{1,\ldots,m\}$ such that $\ker(L - \lambda_i \cdot 1_X) \neq \{0\}$ and thus there exists a $0 \neq w \in \ker(L - \lambda_i \cdot 1_X)$ or in other word there exists a $0 \neq w \in X$ such that $(L - \lambda \cdot 1_X)(w) = 0 \Rightarrow L(w) - \lambda \cdot w = 0 \Rightarrow L(w) = \lambda \cdot w$ proving the theorem. □

Definition 13.52. *Let $\langle X,+,\cdot\rangle$ be a vector space, $Y \subseteq X$ a sub vector space (see 10.108) and $T \in \operatorname{Hom}(X)$ then X is* ***invariant*** *under L if $L(Y) \subseteq Y$ [in other words $\forall u \in Y$ we have $L(u) \in Y$)*

Definition 13.53. *Let $n \in \mathbb{N}$, $\mathcal{F}$ be a field then $M \in \mathcal{M}(n \times n, \mathcal{F})$ (the set of all the n by n square matrices) is a* ***upper triangular matrix*** *if $\forall i,j \in \{1,\ldots,n\}$ with $i > j$ we have $M_{i,j} = 0$*

A special case of a upper triangular matrix is a diagonal matrix.

Definition 13.54. *Let $n \in \mathbb{N}$, $\mathcal{F}$ be a field then $M \in \mathcal{M}(n \times n, \mathcal{F})$ (the set of all the n by n square matrices) is a* ***diagonal matrix*** *if $\forall i,j \in \{1,\ldots,n\}$ with $i \neq j$ we have $M_{i,j} = 0$*

The following theorem gives equivalent conditions for the matrix of a linear operator to be triangular

Theorem 13.55. *Let $\langle X,+,\cdot\rangle$ be a non trivial vector space (hence $n = \dim(V) \in \mathbb{N}$), $\{e_i\}_{i \in \{1,\ldots,n\}}$ a basis for X, $L \in \operatorname{Hom}(X)$ a linear operator the following are equivalent*

1. *$\mathcal{M}(L, \{e_i\}_{i \in \{1,\ldots,n\}}, \{e_i\}_{i \in \{1,\ldots,n\}})$ is upper triangular*
2. *$\forall i \in \{1,\ldots,n\}$ we have $L(e_i) \in \mathcal{S}(\{e_k | k \in \{1,\ldots,i\}\})$ (here $\mathcal{S}$ is the span of $\{e_k\}_{k \in \{1,\ldots,i\}}$ see 10.128)*
3. *$\forall i \in \{1,\ldots,n\}$ we have that $\mathcal{S}(\{e_k | k \in \{1,\ldots,i\}\})$ is invariant under L*

Proof. Let $M = \mathcal{M}(L, \{e_i\}_{i \in \{1,\ldots,n\}}, \{e_i\}_{i \in \{1,\ldots,n\}})$

1 ⇒ 2. Let $i \in \{1,\ldots,n\}$ then we have $L(e_i) = \sum_{k=1}^{n} M_{k,i} \cdot e_k = \sum_{k \in \{1,\ldots,i\}} M_{k,i} \cdot e_k + \sum_{k \in \{i+1,\ldots,n\}} M_{k,i} \cdot e_k = \sum_{k \in \{1,\ldots,i\}} M_{k,i} \cdot e_k \in \mathcal{S}(\{e_k | k \in \{1,\ldots,i\}\})$

2 ⇒ 1. As $\forall i\in\{1,\ldots,n\}$ we have that $L(e_i)\in\mathcal{S}(\{e_k|k\in\{1,\ldots,i\}\})$ we have that $L(e_i)=\sum_{k=1}^{i} L_{k,i}\cdot e_k=\sum_{k=1}^{n} M_{k,i}e_k$ where $M_{k,i}=\left\{\begin{smallmatrix} 0 \text{ if } k>i \\ L_{k,i} \text{ if } k\in\{1,\ldots,i\}\end{smallmatrix}\right.$ so using the definition of $\mathcal{M}(L,\{e_k\}_{k\in\{1,\ldots,n\}},\{e_k\}_{k\in\{1,\ldots,n\}})$ (see 10.300) we have that $\mathcal{M}(L,\{e_j\}_{j\in\{1,\ldots,n\}},\{e_j\}_{j\in\{1,\ldots,n\}})_{k,i}=M_{k,i}$ hence $\mathcal{M}(L,\{e_j\}_{j\in\{1,\ldots,n\}},\{e_j\}_{j\in\{1,\ldots,n\}})_{k,i}=0$ if $k>i$ proving upper triangularity.

2 ⇒ 3. Given $i\in\{1,\ldots,n\}$ take $x\in\mathcal{S}(\{e_k|k\in\{1,\ldots,i\}\})\Rightarrow x=\sum_{k=1}^{i} x_k\cdot e_k$ then as $\forall l\in\{1,\ldots,k\}$ we have $\{e_k|k\in\{1,\ldots,l\}\}\subseteq\{e_k|k\in\{1,\ldots,i\}\}\underset{10.136}{\Rightarrow}\mathcal{S}(\{e_k|k\in\{1,\ldots,l\}\})\subseteq\mathcal{S}(\{e_k|k\in\{1,\ldots,i\}\})$ hence $\forall l\in\{1,\ldots,i\}$ we have $L(e_l)\in\mathcal{S}(\{e_k|k\in\{1,\ldots,i\}\})$. So $L(x)=L(\sum_{k=1}^{i} x_k\cdot e_k)=\sum_{k=1}^{i} x_k\cdot L(e_i)\in\mathcal{S}(\mathcal{S}(\{e_k|k\in\{1,\ldots,i\}\}))\underset{10.135}{=}\mathcal{S}(\{e_k|k\in\{1,\ldots,i\}\})$ proving that $\mathcal{S}(\{e_k|k\in\{1,\ldots,i\}\})$ is invariant under L.

3 ⇒ 2. Given $i\in\{1,\ldots,n\}$ we have as $\mathcal{S}(\{e_k|k\in\{1,\ldots,i\}\})$ is invariant under L that $L(e_i)\in\mathcal{S}(\{e_k|k\in\{1,\ldots,i\}\})$. □

Theorem 13.56. *Let $\langle X,+,\cdot\rangle$ be a complex finite dimensional non trivial vector space with $\dim(X)=n\in\mathbb{N}$ and $L\in\mathrm{Hom}(X)$ then there exists a basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ such that $\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{e_i\}_{i\in\{1,\ldots,n\}})$ is upper triangular.*

Proof. We proof this by induction so let $\mathcal{F}=\{n\in\mathbb{N}|$ if X is a m-dimensional complex vector space with $m\leqslant n$ $L\in\mathrm{Hom}(X)$ then there exists a basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ such that $\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{e_i\}_{i\in\{1,\ldots,n\}})\}$ then

$1\in\mathcal{F}$. If $\dim(X)=1$ then there exists a $\{e_i\}_{i\in\{1,\ldots,1\}}$ such that if $L\in\mathrm{Hom}(X)$ we have $L(e_1)=\sum_{i=1}^{1}\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,1\}},\{e_i\}_{i\in\{1,\ldots,1\}})_{i,1}\cdot e_i=\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,1\}},\{e_i\}_{i\in\{1,\ldots,1\}})$ which as its a 1×1 matrix is upper triangular.

$n\in\mathcal{F}\Rightarrow n+1\in\mathcal{F}$. Assume that $\langle X,+,\cdot\rangle$ is a $n+1$-dimensional complex vector spaces and that $L\in\mathrm{Hom}(X)$ is a linear operator. As L has a eigenvector x with eigenvalue λ (see 13.51) so that $(L-\lambda\cdot I_X)(x)=0$ hence $0\neq x\in\ker(L-\lambda\cdot I_X)\Rightarrow\ker(L-\lambda\cdot I_X)\neq\{0\}$ proving that $\dim(\ker(L-\lambda\cdot I_X))>0$. Hence for $(L-\lambda\cdot I_X)\in\mathrm{Hom}(X)$ we have that $\dim(X)\underset{10.218}{=}\dim(\ker(L-\lambda\cdot I_X))+\mathrm{rank}(L-\lambda\cdot 1_X)$ proving that $\dim((L-\lambda\cdot 1_X)(X))<\dim(X)$. Hence

$$\text{If } U=(L-\lambda\cdot 1_X)(X) \text{ and } \dim(U)<\dim(X) \tag{13.38}$$

Now if $u\in U$ we have $L(u)=L(u)-\lambda\cdot u+\lambda\cdot u=(L-\lambda\cdot 1_X)(u)+\lambda\cdot u\in U$ (as U is a vector space by 10.214) hence

$$L(U)\subseteq U \tag{13.39}$$

So using the above we have that $L_{|U}: U\rightarrow U$ is in $\mathrm{Hom}(U)$ wehere $\dim(U)<\dim(X)=n+1\Rightarrow\dim(U)\leqslant n$ hence as $n\in\mathcal{F}$ we have that $\exists\{u_i\}_{i\in\{1,\ldots,\dim(U)\}}$ a basis for U such that $\mathcal{M}(L_{|U},\{u_i\}_{i\in\{1,\ldots,\dim(U)\}},\{u_i\}_{i\in\{1,\ldots,\dim(U)\}})$ is upper triangular and by 13.55 we have

$$\forall j\in\{1,\ldots,\dim(U)\}\ L(u_j)=(L_{|U})(u_j)\in\mathcal{S}(\{u_k|k\in\{1,\ldots,j\}\}) \tag{13.40}$$

As $\{u_i\}_{i\in\{1,\ldots,\dim(U)\}}$ is linear independent we can extend $\{u_i\}_{i\in\{1,\ldots,\dim(U)\}}$ by 10.203 to a basis $\{w_i\}_{i\in\{1,\ldots,n+1\}}$ of X. Hence using 13.40 we have

$$\forall j\in\{1,\ldots,\dim(U)\}\ L(w_j)\in\mathcal{S}(\{w_k|k\in\{1,\ldots,j\}\}) \tag{13.41}$$

Now $\forall j\in\{\dim(U)+1,\ldots,n+1\}$ we have $L(w_j)=(L-\lambda\cdot 1_X)(w_j)+\lambda\cdot 1_X(w_j)=(L-\lambda\cdot 1_X)(w_j)+\lambda\cdot w_j$. As $(L-\lambda\cdot 1_X)(w_j)\in\operatorname{range}(L-\lambda\cdot 1_X)=U=\mathcal{S}(\{u_k|k\in\{1,\ldots,\dim(U)\}\})=\mathcal{S}(\{w_k|k\in\{1,\ldots,\dim(U)\}\})\subseteq_{10.136}\mathcal{S}(\{w_k|k\in\{1,\ldots,j\}\})$ and $w_j\in\mathcal{S}(\{w_k|k\in\{1,\ldots,j\}\})$ we have as $\mathcal{S}(\{w_k|k\in\{1,\ldots,j\}\})$ is a vector space (see 10.130) we conclude that $L(w_j)\in\mathcal{S}(\{w_k|k\in\{1,\ldots,j\}\})$. To summerize we have that

$$\forall j\in\{\dim(U)+1,\ldots,n+1\}\ L(w_j)\in\mathcal{S}(\{w_k|k\in\{1,\ldots,j\}\}) \tag{13.42}$$

From 13.41 and 13.42 we conclude that $\forall j\in\{1,\ldots,n+1\}$ we have that $L(w_j)\in\mathcal{S}(\{w_k|k\in\{1,\ldots,j\}\})$ hence using 13.55 we have that $\mathcal{M}(L,\{w_i\}_{i\in\{1,\ldots,n+1\}},\{w_i\}_{i\in\{1,\ldots,n+1\}})$ is upper triangular, proving that $n+1\in\mathcal{F}$ □

Theorem 13.57. (Schur's Theorem) *Let $\langle X,\langle\rangle\rangle$ be a real (complex) finite dimensional inner product space with $\dim(X)=n\in\mathbb{N}$ and $L\in\operatorname{Hom}(X)$ then L has a orthonormal basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ such that $\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{e_i\}_{i\in\{1,\ldots,n\}})$ is upper triangular.*

Proof. Using the previous theorem there exissts a basis $\{f_i\}_{i\in\{1,\ldots,n\}}$ such that $\mathcal{M}(L,\{f_i\}_{i\in\{1,\ldots,n\}},\{f_i\}_{i\in\{1,\ldots,n\}})$ is diagonal and using 13.55 we have that $\forall j\in\{1,\ldots,n\}$ we have that $L(\mathcal{S}\{f_i|i\in\{1,\ldots,n\}\})\subseteq\mathcal{S}\{f_i|i\in\{1,\ldots,n\}\}$. Using the Gram-Schmidt procedure (see 12.116) we find a orthonormal basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ such that $\forall j\in\{1,\ldots,n\}$ we have that $\mathcal{S}(\{f_i|i\in\{1,\ldots,n\}\})=\mathcal{S}(e_i|i\in\{1,\ldots,n\})$ hence $L(\{e_i|i\in\{1,\ldots,n\}\})=L(\{f_i|i\in\{1,\ldots,n\}\})\subseteq\{f_i|i\in\{1,\ldots,n\}\}=\{e_i|i\in\{1,\ldots,n\}\}$ hence by 13.55 we have that $\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,\}},\{e_i\}_{i\in\{1,\ldots,n\}})$ is upper diagonal. □

Definition 13.58. *Let $\langle X,\langle\rangle\rangle$ be a real (complex) finite dimensional inner product space then $L\in\operatorname{Hom}(X)$ is **self-adjoint** iff $L=L^*$ (or using the definition of the ajoint of a operator (see 12.120) we have $\forall x\in X$ that $\langle L(x),y\rangle=\langle x,L^*(y)\rangle=\langle x,L(y)\rangle$)*

Theorem 13.59. *Let $\langle X,\langle\rangle\rangle$ be a real (complex) finite dimensional inner product space with a basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ and $L\in\operatorname{Hom}(X)$ then the following are equivalent*

1. *L is self-adjoint*
2. *$\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}})$ is Hermitian*

Proof.

1 ⇒ 2.

$$\begin{aligned}\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}}) &\underset{L^*=L}{=} \mathcal{M}(L^*,\{e_i\}_{i\in\{1,\ldots,n\}})\\ &\underset{12.129}{=} \mathcal{M}(L^*,\{e_i\}_{i\in\{1,\ldots,n\}})^H\end{aligned}$$

2 ⇒ 1. Take $x=\sum_{i=1}^n x_i\cdot e_i\in X$ then we have for $i\in\{1,\ldots,n\}$ that

$$\begin{array}{rcl}(L(x))_i & = & \sum_{k=1}^n \mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}})_{i,k}\cdot x_k\\ & = & \sum_{k=1}^n \mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}})^H_{i,k}\cdot x_k\\ & \underset{12.129}{=} & \sum_{k=1}^n \mathcal{M}(L^*,\{e_i\}_{i\in\{1,\ldots,n\}})_{i,k}\cdot x_k\\ & = & (L^*(x))_i\end{array}$$

□

Corollary 13.60. *Let $\langle X,\langle\rangle\rangle$ be a real finite dimensional inner product space $L\in\operatorname{Hom}(X)$ such that there exists a basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ so that $\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}})$ is diagonal then L is self-adjoint.*

Proof. This follows from the fact that $\forall i,j\in\{1,\ldots,n\}$ we have $(\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}})^H)_{i,j}=(\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}})^*)_{j,i}\underset{\text{real space}}{=}\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}})_{j,i}=\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}})_{i,i}\cdot\delta_{j,i}=\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}})_{i,j}$ so that $\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}})$ is Hermitian and thus by the above theorem we conclude that L is self-adjoint. □

Example 13.61. Let $\langle X,\langle\rangle\rangle$ be a real (complex) finite dimensional inner product space then $1_X\in\operatorname{Hom}(X)$ is self-adjoint as $\forall x,y\in X$ we have $\langle 1_X(x),y\rangle=\langle x,y\rangle=\langle x,1_X(y)\rangle$

Theorem 13.62. *Let $\langle X,\langle\rangle\rangle$ be a real (complex) finite dimensional inner product space and $L\in\operatorname{Hom}(X)$ a self-adjoint operator then every eigenvalue of L is real.*

Proof. In the real case the theorem is trivial, zo let's consider the complex case. Let $x\in X$ be a eigen vector of L with eigenvalue λ then $x\neq 0$ and $L(x)=\lambda\cdot x$. Hence

$$\begin{array}{rcl}\lambda\cdot\|x\|^2 & = & \lambda\cdot\langle x,x\rangle\\ & = & \langle\lambda\cdot x,x\rangle\\ & = & \langle L(x),x\rangle\\ & \underset{L\text{ is self-adjoint}}{=} & \langle L^*(x),x\rangle\\ & = & \langle x,L(x)\rangle\\ & = & \langle x,\lambda\cdot x\rangle\\ & = & \bar{\lambda}\cdot\langle x,x\rangle\\ & = & \bar{\lambda}\cdot\|x\|^2\end{array}$$

Hence as $\|x\|^2\neq 0$ we have $\lambda\in\mathbb{R}$. □

Theorem 13.63. *Let $\langle X,\langle\rangle\rangle$ be a real finite dimensional inner product space and $U\in\operatorname{Hom}(X)$ a unitary operator then* $\det(U^*)=\det(U)=\pm 1$

Proof. We have $1\underset{10.280}{=}\det(1_X)=\det(U^*\circ U)\underset{10.280}{=}\det(U^*)\cdot\det(U)\underset{12.130}{=}\det(U)^2=\det(U^*)^2$ proving that $\det(U)=\pm 1$ and $\det(U^*)=\pm 1$ □

Definition 13.64. (Unitary operator) *Let $\langle X, \langle\rangle\rangle$ be a real (complex) finite dimensional inner product space then $L \in \operatorname{Hom}(X)$ is **unitary** if $L^* \circ L = 1_X$.*

Every unitary operator has a unitary matrix in every basis as is expressed in the following theorem.

Theorem 13.65. *Let $n \in \mathbb{N}$, $\langle X, \langle\rangle\rangle$ be a real (complex) finite dimensional inner product space, $\{e_i\}_{i\in\{1,\ldots,n\}}$ a basis in X and $U \in \operatorname{Hom}(X)$ then we have the following equivalences*

1. *U is unitary*
2. *$\mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}})$ is a unitary matrix*

Proof.

1 ⇒ 2.

$$\begin{aligned}
\mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}})^H \cdot \mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}}) &\underset{12.129}{=} \mathcal{M}(U^*, \{e_i\}_{i\in\{1,\ldots,n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}}) \\
&\underset{10.305}{=} \mathcal{M}(U^* \circ U, \{e_i\}_{i\in\{1,\ldots,n\}}) \\
&= \mathcal{M}(1_X, \{e_i\}_{i\in\{1,\ldots,n\}}) \\
&= E \\
&= \mathcal{M}(1_X, \{e_i\}_{i\in\{1,\ldots,n\}}) \\
&= \mathcal{M}(U \circ U^*, \{e_i\}_{i\in\{1,\ldots,n\}}) \\
&\underset{10.305}{=} \mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}}) \cdot \mathcal{M}(U^*, \{e_i\}_{i\in\{1,\ldots,n\}}) \\
&\underset{12.129}{=} \mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}})^H
\end{aligned}$$

2 ⇒ 1. *Let $x \in X$ then we have that $x = \sum_{i=1}^n x_i \cdot e_i$ and*

$$\begin{aligned}
((U^* \circ U)(x))_i &= \sum_{k=1}^n \mathcal{M}(U^* \circ U, \{e_i\}_{i\in\{1,\ldots,n\}})_{i,k} \cdot x_k \\
&\underset{10.305}{=} \sum_{k=1}^n (\mathcal{M}(U^*, \{e_i\}_{i\in\{1,\ldots,n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}}))_{i,k} \cdot x_k \\
&\underset{12.129}{=} \sum_{k=1}^n (\mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}})^H \cdot \mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}}))_{i,k} \cdot x_k \\
&= \sum_{k=1}^n E_{i,k} \cdot x_k \\
&= \sum_{k=1}^n \delta_{i,k} \cdot x_k \\
&= x_i
\end{aligned}$$

proving that $U^* \circ U = E$

□

a unary operator then $\mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}})$ is a unitary matrix

Proof.

$$\begin{aligned}
\mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}})^H \cdot \mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}}) &\underset{12.129}{=} \mathcal{M}(U^*, \{e_i\}_{i\in\{1,\ldots,n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}}) \\
&\underset{10.305}{=} \mathcal{M}(U^* \circ U, \{e_i\}_{i\in\{1,\ldots,n\}}) \\
&= \mathcal{M}(1_X, \{e_i\}_{i\in\{1,\ldots,n\}}) \\
&= E \\
&= \mathcal{M}(1_X, \{e_i\}_{i\in\{1,\ldots,n\}}) \\
&= \mathcal{M}(U \circ U^*, \{e_i\}_{i\in\{1,\ldots,n\}}) \\
&\underset{10.305}{=} \mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}}) \cdot \mathcal{M}(U^*, \{e_i\}_{i\in\{1,\ldots,n\}}) \\
&\underset{12.129}{=} \mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}})^H
\end{aligned}$$

□

Example 13.66. Let $\langle X, \langle\rangle\rangle$ be a real (complex) finite dimensional inner product space then $1_X \in \mathrm{Hom}(X)$ is unitary.

Proof. $1_X \circ (1_X)^* = (1_X)^* \underset{13.61}{=} 1_X$ □

Theorem 13.67. *Let $\langle X, \langle\rangle\rangle$ be a real (complex) finite dimensional inner product space then the following are equivalent for $L \in \mathrm{Hom}(X)$*

1. *L is unitary*
2. *$\forall x, y \in X$ we have $\langle L(x), L(y)\rangle = \langle x, y\rangle$*

Proof.

1 ⇒ 2. *As L is unitary we have $\forall x, y \in X$ that $\langle x, y\rangle = == \langle L^*(L(x)), y\rangle$*

$$\begin{aligned}
\langle x, y\rangle &= \langle 1_X(x), y\rangle \\
&= \langle (L^* \circ L)(x), y\rangle \\
&= \langle L^*(L(x)), y\rangle \\
&= \langle L(x), (L^*)^*(y)\rangle \\
&\underset{12.121}{=} \langle L(x), L(y)\rangle
\end{aligned}$$

2 ⇒ 1. *We have $\forall x, y \in X$ that*

$$\begin{aligned}
\langle x, (L^* \circ L)(y)\rangle &= \langle x, L^\star(L\, y)\rangle \\
&= \langle L(x), L(y)\rangle \\
&= \langle x, y\rangle \\
&= \langle x, 1_X(y)\rangle
\end{aligned}$$

Hence using 12.97 we have that $\forall y \in X$ we have that $(L^* \circ L)(y) = 1_X(y)$ so that $L^* \circ L = 1_X$. □

Example 13.68. Let $\langle X, \langle\rangle\rangle$ be a real (complex) finite dimensional inner product space then $I: X \to X$ defined by $x \to -x = (-1) \cdot x$ is a unitary operator.

Proof. First as $I(\alpha \cdot x + \beta \cdot y) = -(\alpha \cdot x + \beta \cdot y) = \alpha \cdot (-x) + \beta \cdot (-y) = \alpha \cdot I(x) + \beta \cdot I(x)$ we have that $I \in \operatorname{Hom}(X)$. Further $\langle I(x), I(x)\rangle = \langle(-1) \cdot x, (-1) \cdot x\rangle = (-1) \cdot (-1) \cdot \langle x, x\rangle = \langle x, x\rangle$ □

Theorem 13.69. *Let $\langle X, \langle\rangle\rangle$ be a real (complex) finite dimensional inner product space and $L \in \operatorname{Hom}(X)$ a unitary operator then we have*

1. $L \circ L^* = 1_X$
2. $L^* = L^{-1}$, $L = (L^*)^{-1}$ *and* L^*, L^{-1} *are unitary*
3. L *is self-adjoint*
4. L *is a linear isometry (in the normed space generated by the inner product)*

Proof.

1. $L^* \circ L \underset{12.123}{=} (L \circ L^*)^* = (1_X)^* \underset{13.61}{=} 1_X$
2. From the definition and (1) we have that $L^* \circ L = 1_X = L \circ L^*$ proving that L^* is unitary and $L^* = L^{-1}$, $L = (L^*)^{-1}$.
3. This is trivial as in (2) we saw that $L^* \circ L = 1_X = L \circ L^*$
4. Using (2) we have that L is a bijection and $\forall x \in X$ we have $\|L(x)\| = \sqrt{\langle L(x), L(x)\rangle} \underset{13.67}{=} \sqrt{\langle x, x\rangle} = \|x\|$ proving that L is a linear isomorphism. □

The following theorem shows that every mapping of one basis to a orthonormal basis is

Theorem 13.70. *Let $\langle X, \langle\rangle\rangle$ be a real (complex) finite dimensional inner product space with two basis $\{e_i\}_{i\in\{1,\ldots,n\}}$, $\{f_i\}_{i\in\{1,\ldots,n\}}$ then for the mappings U, $V \in \operatorname{Hom}(X)$ defined by $\forall i \in \{1, \ldots, n\}$ $U(e_i) = f_i$ and $V(f_i) = e_i$ (U, V are called coordinate trasnformations) then we have that*

1. $U \circ V = 1_X = V \circ U$ *(or equivalently* $U = V^{-1}$ *and* $V = U^{-1}$
2. $\mathcal{M}(V, \{e_i\}_{i\in\{1,\ldots,n\}}) = \mathcal{M}(U_i, \{e_i\}_{i\in\{1,\ldots,n\}})^{-1}$
3. *If* $\{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,n\}}$ *are orthonormal then* U *is unitary (see 12.127)*
4. *If* $\{e_i\}_{i\in\{1,\ldots,n\}}, \{f_i\}_{i\in\{1,\ldots,n\}}$ *are orthonormal then* $\mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}})^{-1} = \mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}})^H$ *(proving that* $\mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}})$ *is unitary*
5. *If* $L \in \operatorname{Hom}(X)$ *then we have that*

   $\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}) = \mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}}) \cdot \mathcal{M}(L, \{f_i\}_{i\in\{1,\ldots,n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}})^{-1}$

   or bringing over the matrices and applying the inverses

   $\mathcal{M}(L, \{f_i\}_{i\in\{1,\ldots,n\}}) = \mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}})^{-1} \cdot \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i\in\{1,\ldots,n\}})$

Proof.

1. Let $x \in X$ then $x = \sum_{i=1}^{n} x_i \cdot e_i = \sum_{i=1}^{n} x'_i \cdot f_i$ and we have that

$$\begin{aligned}(U \circ V)(x) &= \left(U\left(V\left(\sum_{i=1}^{n} x'_i \cdot f_i\right)\right)\right)\\ &= \left(U\left(\sum_{i=1}^{n} x'_i \cdot V(f_i)\right)\right)\\ &= U\left(\sum_{i=1}^{n} x'_i \cdot e_i\right)\\ &= \sum_{i=1}^{n} x'_i \cdot U(e_i)\\ &= \sum_{i=1}^{n} x'_i \cdot f_i\\ &= x\\ &= 1_X(x)\end{aligned}$$

proving that $U \circ V = 1_X$. Like wise we have

$$\begin{aligned}(V \circ U)(x) &= \left(V\left(U\left(\sum_{i=1}^{n} x_i \cdot e_i\right)\right)\right)\\ &= \left(V\left(\sum_{i=1}^{n} x_i \cdot U(e_i)\right)\right)\\ &= V\left(\sum_{i=1}^{n} x_i \cdot f_i\right)\\ &= \sum_{i=1}^{n} x_i \cdot V(f_i)\\ &= \sum_{i=1}^{n} x_i \cdot e_i\\ &= x\\ &= 1_X(x)\end{aligned}$$

proving that $V \circ U = 1_X$

2. This follows from the fact that

$$\begin{aligned}\mathcal{M}(U, \{e_i\}_{i \in \{1,\ldots,n\}}) \cdot \mathcal{M}(V, \{e_i\}_{i \in \{1,\ldots,n\}}) &\underset{10.305}{=} \mathcal{M}(U \circ V, \{1, \ldots, n\})\\ &\underset{(1)}{=} \mathcal{M}(1_X, \{1, \ldots, n\})\\ &= E\\ &= \mathcal{M}(1_X, \{1, \ldots, n\})\end{aligned}$$

$$\underset{(1)}{=} \mathcal{M}(V \circ U, \{1, ..., n\})$$
$$\underset{10.305}{=} \mathcal{M}(V, \{e_i\}_{i \in \{1,...,n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1,...,n\}})$$

3. Let $x, y \in X$ then $x = \sum_{i=1}^n x_i \cdot e_i = \sum_{i=1}^n x_i' \cdot f_i$ and $y = \sum_{i=1}^n y_i \cdot e_i = \sum_{i=1}^n y_i' \cdot f_i$ then we have

$$\begin{aligned}
\langle U(x), U(y)\rangle &= \left\langle U\left(\sum_{i=1}^n x_i \cdot e_i\right), U\left(\sum_{k=1}^n y_k \cdot e_k\right)\right\rangle \\
&= \left\langle \sum_{i=1}^n x_i \cdot U(e_i), \left(\sum_{k=1}^n y_k \cdot U(e_k)\right)\right\rangle \\
&= \left\langle \sum_{i=1}^n x_i \cdot f_i, \left(\sum_{k=1}^n y_k \cdot f_k\right)\right\rangle \\
&= \sum_{i=1}^n x_i \cdot \left(\left\langle f_i, \sum_{k=1}^n y_k \cdot f_k\right\rangle\right) \\
&= \sum_{i=1}^n \left(\sum_{k=1}^n x_i \cdot \bar{y_i} \cdot \langle f_i, f_k\rangle\right) \\
&\underset{\text{orthornomality}\{f_i\}_{i \in \{1,...,n\}}}{=} \sum_{i=1}^n \left(\sum_{k=1}^n x_i \cdot \bar{y_i} \cdot \delta_{i,k}\right) \\
&\underset{\text{orthonormality of}\{e_i\}_{i \in \{1,...,n\}}}{=} \sum_{i=1}^n \left(\sum_{k=1}^n x_i \cdot \bar{y_i} \cdot \langle e_i, e_k\rangle\right) \\
&= \sum_{i=1}^n x_i \cdot \left(\sum_{k=1}^n \langle e_i, y_k \cdot e_k\rangle\right) \\
&= \sum_{i=1}^n x_i \cdot \left\langle e_i, \sum_{k=1}^n y_k \cdot e_k\right\rangle \\
&= \left\langle \sum_{i=1}^n x_i \cdot e_i, \sum_{k=1}^n y_k \cdot e_k\right\rangle \\
&= \langle x, y\rangle
\end{aligned}$$

Using 13.67 we have then that U is a unitary operator.

4. This follows from

$$\begin{aligned}
\mathcal{M}(U, \{e_i\}_{i \in \{1,...,n\}})^H \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1,...,n\}}) &\underset{10.338}{=} \mathcal{M}(U^*, \{e_i\}_{i \in \{1,...,n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1,...,n\}}) \\
&\underset{10.305}{=} \mathcal{M}(U^* \circ U, \{1, ..., n\}) \\
&\underset{(1)}{=} \mathcal{M}(1_X, \{1, ..., n\}) \\
&= E
\end{aligned}$$

$$
\begin{aligned}
&= \mathcal{M}(1_X, \{1, \ldots, n\}) \\
&\underset{(1)}{=} \mathcal{M}(U \circ U^*, \{1, \ldots, n\}) \\
&= \mathcal{M}(U, \{e_i\}_{i \in \{1, \ldots, n\}}) \cdot \mathcal{M}(U^*, \{e_i\}_{i \in \{1, \ldots, n\}}) \\
&\underset{10.338}{=} \mathcal{M}(U, \{e_i\}_{i \in \{1, \ldots, n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1, \ldots, n\}})^H
\end{aligned}
$$

5. Given $i \in \{1, \ldots, n\}$ we have that

$$
\begin{aligned}
L(e_i) &= L(V(f_i)) \\
&= L\left(\sum_{k=1}^{n} \mathcal{M}(V, \{f_i\}_{i \in \{1, \ldots, n\}})_{k,j} \cdot f_k\right) \\
&= \sum_{k=1}^{n} \mathcal{M}(V, \{f_i\}_{i \in \{1, \ldots, n\}})_{k,j} \cdot L(f_k) \\
&= \sum_{k=1}^{n} \left(\mathcal{M}(V, \{f_i\}_{i \in \{1, \ldots, n\}})_{k,j} \cdot \sum_{l=1}^{n} \mathcal{M}(L, \{f_i\}_{i \in \{1, \ldots, n\}})_{l,k} \cdot f_l\right) \\
&= \sum_{k=1}^{n} \left(\mathcal{M}(V, \{f_i\}_{i \in \{1, \ldots, n\}})_{k,j} \cdot \sum_{l=1}^{n} \mathcal{M}(L, \{f_i\}_{i \in \{1, \ldots, n\}})_{l,k} \cdot (U(e_l))\right) \\
&= \sum_{k=1}^{n} \left(\mathcal{M}(V, \{f_i\}_{i \in \{1, \ldots, n\}})_{k,j} \cdot \sum_{l=1}^{n} \mathcal{M}(L, \{f_i\}_{i \in \{1, \ldots, n\}})_{l,k} \cdot \left(\sum_{r=1}^{n} \mathcal{M}(U, \{e_i\}_{i \in \{1, \ldots n\}})_{r,l} \cdot e_r\right)\right) \\
&= \sum_{k=1}^{n} \left(\mathcal{M}(V, \{f_i\}_{i \in \{1, \ldots, n\}})_{k,j} \cdot \sum_{l=1}^{n} \left(\sum_{r=1}^{n} \mathcal{M}(L, \{f_i\}_{i \in \{1, \ldots, n\}})_{l,k} \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1, \ldots n\}})_{r,l} \cdot e_r\right)\right) \\
&\underset{10.48}{=} \sum_{k=1}^{n} \left(\mathcal{M}(V, \{f_i\}_{i \in \{1, \ldots, n\}})_{k,j} \cdot \sum_{r=1}^{n} \left(\sum_{l=1}^{n} \mathcal{M}(L, \{f_i\}_{i \in \{1, \ldots, n\}})_{l,k} \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1, \ldots n\}})_{r,l} \cdot e_r\right)\right) \\
&= \sum_{k=1}^{n} \left(\mathcal{M}(V, \{f_i\}_{i \in \{1, \ldots, n\}})_{k,j} \cdot \sum_{r=1}^{n} (\mathcal{M}(U, \{e_i\}_{i \in \{1, \ldots, n\}}) \cdot \mathcal{M}(L, \{f_i\}_{i \in \{1, \ldots, n\}}))_{r,k} \cdot e_r\right)
\end{aligned}
$$

$$\begin{aligned}
&= \sum_{k=1}^{n}\left(\sum_{r=1}^{n}\mathcal{M}(V,\{f_i\}_{i\in\{1,\ldots,n\}})_{k,j}\cdot(\mathcal{M}(U,\{e_i\}_{i\in\{1,\ldots,n\}})\cdot\mathcal{M}(L,\{f_i\}_{i\in\{1,\ldots,n\}}))_{r,k}\cdot e_r\right)\\
&\underset{10.48}{=} \sum_{r=1}^{n}\left(\sum_{k=1}^{n}\mathcal{M}(V,\{f_i\}_{i\in\{1,\ldots,n\}})_{k,j}\cdot(\mathcal{M}(U,\{e_i\}_{i\in\{1,\ldots,n\}})\cdot\mathcal{M}(L,\{f_i\}_{i\in\{1,\ldots,n\}}))_{r,k}\cdot e_r\right)\\
&= \sum_{r=1}^{n}(\mathcal{M}(U,\{e_i\}_{i\in\{1,\ldots,n\}})\cdot\mathcal{M}(L,\{f_i\}_{i\in\{1,\ldots,n\}})\cdot\mathcal{M}(V,\{f_i\}_{i\in\{1,\ldots,n\}}))_{r,j}\cdot e_r\\
&\underset{10.312}{=} \sum_{r=1}^{n}(\mathcal{M}(U,\{e_i\}_{i\in\{1,\ldots,n\}})\cdot\mathcal{M}(L,\{f_i\}_{i\in\{1,\ldots,n\}})\cdot\mathcal{M}(U,\{\mathrm{fe}_i\}_{i\in\{1,\ldots,n\}})^{-1})_{r,j}\cdot e_r
\end{aligned}$$

hence we have by definition of the matrix of a linear function proved that

$$\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}})=\mathcal{M}(U,\{e_i\}_{i\in\{1,\ldots,n\}})\cdot\mathcal{M}(L,\{f_i\}_{i\in\{1,\ldots,n\}})\cdot\mathcal{M}(U,\{e_i\}_{i\in\{1,\ldots,n\}})^{-1} \tag{13.43}$$ □

Definition 13.71. (Normal operator) *Let $\langle X,\langle\rangle\rangle$ be a real (complex) finite dimensional inner product space then $L\in\mathrm{Hom}(X)$ is **normal** if $L\circ L^*=L^*\circ L$*

Lemma 13.72. *Let $\langle X,\langle\rangle\rangle$ be a real (complex) finite dimensional inner product space then every self-adjoint operator is normal. Hence every unitary operator is normal (see 13.69).*

Proof. This is trivial as $L\circ L^*=L\circ L=L^*\circ L$ □

Theorem 13.73. *Let $\langle X,\langle\rangle\rangle$ be a complex inner product vector space, $L\in\mathrm{Hom}(X)$ then if $\forall x\in X$ we have $\langle L(x),x\rangle=0$ we have $L=0$*

Proof. Let $x,y\in X$ then we have

$$\begin{aligned}
\langle L(x+y),x+y\rangle-\langle L(x-y),x-y\rangle &= \langle L(x)+L(y),x+y\rangle-\langle L(x)-L(y),x-y\rangle\\
&= \langle L(x),x\rangle+\langle L(x),y\rangle+\langle L(y),x\rangle+\langle L(y),y\rangle-\langle L(x),x\rangle+\langle L(x),y\rangle+\langle L(y),x\rangle-\langle L(y),y\rangle\\
&= 2\cdot\langle L(x),y\rangle+2\cdot\langle L(y),x\rangle
\end{aligned}$$

and

$$\begin{aligned}\langle L(x+i\cdot y), x+i\cdot y\rangle - \langle L(x-i\cdot y),\\ x-i\cdot y\rangle &= \langle L(x)+i\cdot L(y), x+i\cdot y\rangle - \langle L(x)-i\cdot L(y), x-i\cdot y\rangle\\ &= \langle L(x),x\rangle + \langle L(x), i\cdot y\rangle + \langle i\cdot L(y), x\rangle + \langle i\cdot L(y), i\cdot y\rangle - \langle L(x),x\rangle + \langle L(x), i\cdot y\rangle + \langle i\cdot L(y), x\rangle - \langle i\cdot L(y), i\cdot y\rangle\\ &= \langle L(x), i\cdot y\rangle + \langle i\cdot L(y), x\rangle + \langle L(x), i\cdot y\rangle + \langle i\cdot L(y), x\rangle\\ &= -2\cdot i\cdot\langle L(x), y\rangle + 2\cdot i\cdot\langle L(y), x\rangle\end{aligned}$$

Hence we have

$$\begin{aligned}(\langle L(x+y), x+y\rangle - \langle L(x-y), x-y\rangle) +\\ i\cdot(\langle L(x+i\cdot y), x+i\cdot y\rangle - \langle L(x-i\cdot y),\\ x-i\cdot y\rangle) &= 2\cdot\langle L(x), y\rangle + 2\cdot\langle L(y), x\rangle + 2\cdot\langle L(x), y\rangle - 2\cdot\langle L(y), x\rangle\\ &= 4\cdot\langle L(x), y\rangle\end{aligned}$$

proving that

$$\begin{aligned}\langle L(x), y\rangle &= \frac{\langle L(x+i\cdot y), x+i\cdot y\rangle - \langle L(x-i\cdot y), x-i\cdot y\rangle}{4} + i\cdot\frac{\langle L(x+i\cdot y), x+i\cdot y\rangle - \langle L(x-i\cdot y), x-i\cdot y\rangle}{4}\\ &= \frac{0-0}{4} + i\cdot\frac{0-0}{4}\\ &= 0\end{aligned}$$

Hence if we take $y = L(x)$ we have that $\|L(x)\| = \langle L(x), L(x)\rangle = 0$ so that $L(x) = 0$ and as x is choosen arbitray we have $L = 0$. □

Theorem 13.74. *Let $\langle X, \langle\rangle\rangle$ be a real (complex) finite dimensional inner product space and $L \in \mathrm{Hom}(X)$ a self-adjoint operator then $L = 0$ if and only if $\forall x \in X$ we have $\langle L(x), x\rangle = 0$*

Proof. If $\langle X, \langle\rangle\rangle$ is a complex space the theorem is proved by the previous theorem (see 13.73) (even with asumption of self-adjointness) so we must only prove the case of a real inner product space. Let $x, y \in X$ then

$$\begin{aligned}\langle L(x+y), x+y\rangle - \langle L(x-y), x-y\rangle &= \langle L(x)+L(y), x+y\rangle - \langle L(x)-L(y), x-y\rangle\\ &= \langle L(x), x\rangle + \langle L(x), y\rangle + \langle L(y), x\rangle + \langle L(y), y\rangle - (\langle L(x), x\rangle - \langle L(x), y\rangle - \langle L(y), x\rangle + \langle L(y), y\rangle)\\ &= \langle L(x), x\rangle + \langle L(x), y\rangle + \langle L(y), x\rangle + \langle L(y), y\rangle - \langle L(x), x\rangle + \langle L(x), y\rangle + \langle L(y), x\rangle - \langle L(y), y\rangle\end{aligned}$$

$$\begin{aligned}
&= \langle L(x), y\rangle + \langle L(y), x\rangle + \langle L(x), y\rangle + \\
&\quad \langle L(y), x\rangle \\
&= 2\cdot\langle L(x), y\rangle + 2\cdot\langle L(y), x\rangle \\
&\underset{X\,\text{is real}}{=} 2\cdot\langle L(x), y\rangle + 2\cdot\langle x, L(y)\rangle \\
&= 2\cdot\langle L(x), y\rangle + 2\cdot\langle L^*(x), y\rangle \\
&\underset{L=L^*}{=} 2\cdot\langle L(x), y\rangle + 2\cdot\langle L(x), y\rangle \\
&= 4\cdot\langle L(x), y\rangle
\end{aligned}$$

hence $\langle L(x), y\rangle = \frac{\langle L(x+y), x+y\rangle - \langle L(x-y), x-y\rangle}{4} = \frac{0-0}{4} = 0$. So taking $y = L(x)$ we have that $\|L(x)\| = \langle L(x), L(x)\rangle = 0$ proving as x is choosen arbitrary that $L = 0$. □

Lemma 13.75. *Let $\langle X, \langle\rangle\rangle$ be a real (complex) finite dimensional inner product space together with the inner product norm $\|\|$ (see 12.102) then $L \in \mathrm{Hom}(X)$ is normal if and only if $\forall x \in X$ we have $\|L(x)\| = \|L^*(x)\|$*

Proof. First we have

$$\begin{aligned}
(L\circ L^* + L^*\circ L)^* &\underset{12.122}{=} (L\circ L^*)^* + (L^*\circ L)^* \\
&\underset{12.123}{=} L^*\circ(L^*)^* + (L^*)^*\circ L^* \\
&\underset{12.121}{=} L^*\circ L + L\circ L^* \\
&= L\circ L^* + L^*\circ L
\end{aligned}$$

proving that

$$L\circ L^* + L^*\circ L\,\text{is self adjoint} \tag{13.44}$$

$$\begin{aligned}
L\,\text{is normal} &\Leftrightarrow L\circ L^* = L^*\circ L \\
&\Leftrightarrow L\circ L^\star - L^*\circ L = 0 \\
&\underset{13.44\ \text{and}\ 13.74}{\Leftrightarrow} \forall x\in X \vDash \langle (L\circ L^* - L^*\circ L)(x), x\rangle = 0 \\
&\Leftrightarrow \forall x\in X \vDash \langle (L\circ L^*(x) x)\rangle = \langle (L^*\circ L)(x), x\rangle \\
&\Leftrightarrow \forall x\in X \vDash \langle L(L^*(x)), x\rangle = \langle L^*(L(x)), x\rangle \\
&\Leftrightarrow \forall x\in X \vDash \langle L^*(x), L^*(x)\rangle = \langle L(x), (L^*)^*(x)\rangle \\
&\underset{12.121}{\Leftrightarrow} \forall x\in X \vDash \langle L^*(x), L^*(x)\rangle = \langle L(x), L(x)\rangle \\
&\Leftrightarrow \forall x\in X \vDash \|L^*(x)\| = \|L(x)\|
\end{aligned}$$

proving the theorem. □

Theorem 13.76. *Let $\langle X, \langle\rangle\rangle$ be a real (complex) finite dimensional inner product space with $\dim(X) = n \in \mathbb{N}$ and $L \in \mathrm{Hom}(X)$ then the following are equivalent*

1. *There exists a orthonormal basis $\{e\}_{i\in\{1,\ldots,n\}}$ of X such that $\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})$ is diagonal*
2. *X has a orthonormal basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ where $\forall i\in\{1,\ldots,n\}$ e_i is a eigen vector of L*

Proof.

1 ⇒ 2. *As there exists a basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ such that $M=\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{e_i\}_{i\in\{1,\ldots,n\}})$ is diagonal we have that $\forall i\in[1,\ldots,n]$ that $L(e_i)=\sum_{k=1}^n M_{k,i}\cdot e_k=M_{i,i}\cdot e_i\Rightarrow e_i$ is a eignvector with eigen value $M_{i,i}$.*

2 ⇒ 1. *As $\{e_i\}_{i\in\{1,\ldots,n\}}$ is a set of eigen vectors with eigen values λ_i we have if we define M by $M_{i,j}=\delta_{i,j}\cdot\lambda_i$ so that $L(e_i)=\lambda_i\cdot e_i=\sum_{k=1}^n\delta_{k,i}\cdot\lambda_i\cdot e_k=\sum_{k=1}^n M_{k,j}\cdot e_k$ proving that $\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{e_i\}_{i\in\{1,\ldots,n\}})=M$ which is by construction diagonal.* □

Theorem 13.77. (Complex Spectral Theorem) *Let $\langle X,\langle\rangle\rangle$ be a complex finite dimensional inner product space with $\dim(X)=n\in\mathbb{N}$ and $L\in\mathrm{Hom}(X)$ then the following are equivalent*

1. *L is normal*
2. *There exists a orthonormal basis $\{e\}_{i\in\{1,\ldots,n\}}$ of X such that $\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{e_i\}_{i\in\{1,\ldots,n\}})$ is diagonal.*
3. *X has a orthonormal basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ where $\forall i\in\{1,\ldots,n\}$ e_i is a eigen vector of L*

Proof.

1 ⇒ 2. As L is normal we have using Schur's theorem (see 13.57) there exists a orthormal basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ of X such that $M=\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}},\{e_i\}_{i\in\{1,\ldots,n\}})$ is a upper triangluar matrix. Define now $\mathcal{S}=\{m\in\mathbb{N}|\text{If } m\leqslant n \text{ then } \forall i\in\{1,\ldots,m\}\wedge\forall j\in\{i+1,\ldots,n\} \text{ we have that } M_{i,j}=0\}$ then we have

$1\in\mathcal{S}$. As M is upper triangular we have that $L(e_1)=\sum_{k=1}^n M_{k,1}\cdot e_k=M_{1,1}\cdot e_1$ so that $\|L(e_1)\|^2=\langle M_{1,1}\cdot e_1,M_{1,1}\cdot e_1\rangle=M_{1,1}\cdot\overline{M_{1,1}}\cdot\langle e_1,e_1\rangle=|M_{1,1}|^2$ proving that

$$\|L(e_1)\|^2=|M_{1,1}|^2 \tag{13.45}$$

Further if $M^*=\mathcal{M}(L^*,\{e_i\}_{i\in\{1,\ldots,n\}},\{e_i\}_{i\in\{1,\ldots,n\}})$ then using 12.129 we have that $M_{i,j}=\overline{M_{j,i}}$ so that

$$\begin{aligned}
\|L^*(e_1)\|^2 &= \langle L(e_1),L(e_1)\rangle\\
&= \left\langle\sum_{k=1}^n M^*_{k,1}\cdot e_k,\sum_{l=1}^n M^*_{l,1}\cdot e_l\right\rangle\\
&= \sum_{k=1}^n M^*_{k,1}\cdot\left\langle e_k,\sum_{l=1}^n M^*_{k,1}\cdot e_l\right\rangle\\
&= \sum_{k=1}^n M^\star_{k,1}\cdot\left(\sum_{l=1}^n M^*_{k,1}\cdot\langle e_k,e_l\rangle\right)\\
&= \sum_{k=1}^n M^\star_{k,1}\cdot\left(\sum_{l=1}^n M^*_{k,1}\cdot\delta_{k,l}\right)
\end{aligned}$$

$$= \sum_{k=1}^{n} M_{k,1}^{\star} \cdot \overline{M_{k,1}^{*}}$$
$$= \sum_{k=1}^{n} \overline{M_{1,k}} \cdot \overline{\overline{M_{1,k}}}$$
$$= \sum_{k=1}^{n} \overline{M_{1,k}} \cdot M_{1,k}$$
$$= \sum_{k=1}^{n} |M_{1,k}|^2$$

proving that

$$\|L^*(e_1)\|^2 = \sum_{k=1}^{n} |M_{1,k}|^2 \tag{13.46}$$

As by 13.75 we have $\|L(e_1)\| = \|L^*(e_1)\| \underset{13.45 \text{ and } 13.46}{\Rightarrow} |M_{1,1}|^2 = |M_{1,1}|^2 + \sum_{k=2}^{n} |M_{1,k}$ proving that $0 = \sum_{k=2}^{n} |M_{1,k}|^2$ so that $\forall k \in \{2, \ldots n\}$ we have that $|M_{1,k}| = 0$ or $\forall i \in \{1\}$ we have $\forall j \in \{i+1, \ldots, n\}$ that $M_{1,j} = 0$. Hence $1 \in \mathcal{S}$.

$\boldsymbol{m \in \mathcal{S} \Rightarrow m+1 \in \mathcal{S}}$**.** If n<m+1 then $m+1 \in \mathcal{S}$ so we have to consider the case that $m+1 \leqslant n$. As $m \in \mathcal{S}$ we have

$$\forall i \in \{1, \ldots, m\} \text{ and } \forall j \in \{i+1, \ldots, n\} \text{ we have } M_{i,j} = 0 \tag{13.47}$$

Now $L(e_{m+1}) = \sum_{k=1}^{n} M_{k,m+1} \cdot e_k \underset{M \text{ is upper triangular}}{=} \sum_{k=1}^{m+1} M_{k,m+1} \cdot e_k = M_{m+1,m+1} \cdot e_{m+1} + \sum_{k=1}^{m} M_{k,m+1} \cdot e_k \underset{13.47}{=} M_{m+1,m+1} \cdot e_{m+1}$ hence $\|L(e_{m+1})\|^2 = \langle L(e_{m+1}), L(e_{m+1}) \rangle = \langle M_{m+1,m+1} \cdot e_{m+1}, M_{m+1} \rangle = M_{m+1,m+1} \cdot \overline{M_{m+1,m+1}} \cdot \langle e_{m+1}, e_{m+1} \rangle$. So we have

$$\|L(e_{m+1})\|^2 = |M_{m+1,m+1}|^2 \tag{13.48}$$

Further if $M^* = \mathcal{M}(L^*, \{e_i\}_{i \in \{1, \ldots, n\}}, \{e_i\}_{i \in \{1, \ldots, n\}})$ then using 12.129 we have that $M_{i,j} = \overline{M_{j,i}}$ so that

$$\begin{aligned}
\|L^*(e_{m+1})\|^2 &= \langle L(e_{m+1}), L(e_{m+1}) \rangle \\
&= \left\langle \sum_{k=1}^{n} M_{k,m+1}^{*} \cdot e_k, \sum_{l=1}^{n} M_{l,m+1}^{*} \cdot e_l \right\rangle \\
&= \sum_{k=1}^{n} M_{k,m+1}^{*} \cdot \left\langle e_k, \sum_{l=1}^{n} M_{k,m+1}^{*} \cdot e_l \right\rangle \\
&= \sum_{k=1}^{n} M_{k,m+1}^{\star} \cdot \left(\sum_{l=1}^{n} M_{k,m+1}^{*} \cdot \langle e_k, e_l \rangle \right) \\
&= \sum_{k=1}^{n} M_{k,m+1}^{\star} \cdot \left(\sum_{l=1}^{n} M_{k,m+1}^{*} \cdot \delta_{k,l} \right) \\
&= \sum_{k=1}^{n} M_{k,m+1}^{\star} \cdot \overline{M_{k,m+1}^{*}}
\end{aligned}$$

$$\begin{aligned} &= \sum_{k=1}^{n} \overline{M_{m+1,k}} \cdot \overline{\overline{M_{m+1,k}}} \\ &= \sum_{k=1}^{n} \overline{M_{m+1,k}} \cdot M_{m+1,k} \\ &= \sum_{k=1}^{n} |M_{m+1,k}|^2 \end{aligned}$$

proving that

$$\|L^*(e_{m+1})\|^2 = \sum_{k=1}^{n} |M_{m+1,k}|^2 \tag{13.49}$$

As by 13.75 we have $\|L(e_{m+1})\| = \|L^*(e_{m+1})\| \underset{13.48 \text{ and } 13.49}{\Rightarrow} |M_{m+1,m+1}|^2 = |M_{m+1,m+1}|^2 + \sum_{k\in\{1,\ldots,n\}\setminus\{m+1\}} |M_{m+1,k}|^2$ proving that $0 = \sum_{k\in\{1,\ldots,n\}\setminus\{m+1\}} |M_{m+1,k}|^2$ so that $\forall k \in \{1,\ldots n\}\setminus\{m+1\}$ we have that $|M_{m+1,k}| = 0$. Hence $\forall j \in \{(m+1)+1,\ldots,n\}$ we have $M_{m+1,j} = 0$ and using 13.47 we conclude that $\forall i \in \{1,\ldots,m+1\}$ we have $\forall j \in \{i+1,\ldots,n\}$ we have $M_{i,j} = 0$.

Mathematical induction proves that

$$\mathcal{S} = \mathbb{N} \tag{13.50}$$

Hence if $i, j \in \{1,\ldots,n\}$ with $i \neq j$ we have either

***i < j*.** As $i \in \mathbb{N} = \mathcal{S}$ we have as $i \leqslant n$ and $j \in \{i+1,\ldots,n\}$ that $M_{i,j} = 0$

***j < i*.** As M is upper diagonal we have $M_{i,j} = 0$

proving that $M = \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})$ is diagonal.

2 ⇒ 1. As there exists a basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ such that $M = \mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})$ is diagonal and $M^* = \mathcal{M}(L^\times, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e\})$ is defined by $M^\times_{i,j} = \overline{M_{j,i}}$ (see 10.338) we have that M^* is also diagonal. The we have $\forall j \in \{1,\ldots,n\}$ that

$$\begin{aligned} (L \circ L^*)(e_j) &= L(L^*(e_j)) \\ &= L\left(\sum_{k=1}^{n} M^*_{k,j} \cdot e_k\right) \\ &= L(M^\star_{j,j} \cdot e_j) \\ &= M^*_{j,j} \cdot L(e_j) \\ &= M^*_{j,j} \cdot \left(\sum_{k=1}^{n} M_{k,j} \cdot e_k\right) \\ &= M^*_{j,j} \cdot M_{j,j} \cdot e_j \\ &= M_{i,j} \cdot M^*_{i,j} \cdot e_j \\ &= M_{j,j} \cdot \sum_{k=1}^{n} M^*_{k,j} \cdot e_k \end{aligned}$$

$$\begin{aligned}
&= M_{j,j}\cdot L^{\star}(e_j)\\
&= L^{*}(M_{j,j}\cdot e_j)\\
&= L^{*}\left(\sum_{k=1}^{n} M_{k,j}\cdot e_k\right)\\
&= L^{*}(L(e_j))
\end{aligned}$$

proving that

$$\forall j\in\{1,\ldots,n\}\text{ we have }(L\circ L^{*})(e_j)=(L^{*}\circ L)(e_j) \tag{13.51}$$

Hence if $x\in X$ we have that $x=\sum_{i=1}^{n} x_i\cdot e_j$ and thus

$$\begin{aligned}
(L\circ L^{*})(x) &= (L\circ L^{*})\left(\sum_{i=1}^{n} x_i\cdot e_i\right)\\
&= \sum_{i=1}^{n} x_i\cdot (L\circ L^{*})(e_i)\\
&= \sum_{i=1}^{n} x_i\cdot (L^{*}\circ L)(e_i)\\
&= (L^{*}\circ L)\left(\sum_{i=1}^{n} x_i\cdot e_i\right)\\
&= (L^{*}\circ L)(x)
\end{aligned}$$

proving that $L\circ L^{*}=L^{*}\circ L$ and thus that L is normal.

2 ⇔ 3. This follows from 13.76. □

Corollary 13.78. *Let $\langle X,\langle\rangle\rangle$ be a complex finite dimensional inner product space with $\dim(X)=n\in\mathbb{N}$, $\{e_i\}_{i\in\{1,\ldots,n\}}$ a orthonormal base, $L\in\operatorname{Hom}(X)$ a normal operator then there exists a unary operator $U\in\operatorname{Hom}(X)$ such that $\mathcal{M}(U^{*}\circ L\circ U,$ $\{e_i\}_{i\in\{1,\ldots,n\}})$ is diagonal.*

Proof. Using the Complex spectral theorem 13.77 there exists a basis $\{f_i\}_{i\in\{1,\ldots,n\}}$ such that $\mathcal{M}(L,\{f_i\}_{i\in\{1,\ldots,n\}})$ is diagonal. Take now the unitary mapping $U\in\operatorname{Hom}(X)$ defined by $\forall i\in\{1,\ldots,n\}$ $U(e_i)=f_i$ (see 13.70) then we have $\forall i,j\in\{1,\ldots,$ $n\}$ that

$$\begin{aligned}
\mathcal{M}(U^{*}\circ L\circ U,\{e_i\}_{i\in\{1,\ldots,n\}})_{i,j} &\underset{12.115}{=} \langle (U^{*}\circ L\circ U)(e_j),e_i\rangle\\
&= \langle U^{*}(L(U(e_j))),e_i\rangle\\
&= \langle U^{*}(L(f_j)),e_i\rangle\\
&= \langle L(f_j),(U^{*})^{*}(e_i)\rangle\\
&= \langle L(f_j),U(e_i)\rangle\\
&= \langle L(f_j),f_i\rangle\\
&\underset{12.115}{=} \mathcal{M}(L,\{f_i\}_{i\in\{1,\ldots,n\}})_{i,j}\\
&= \mathcal{M}(L,\{f_i\}_{i\in\{1,\ldots,n\}})_{i,i}\cdot\delta_{i,j}
\end{aligned}$$

□

We need the following theorem to prove the spectral theorem in the real case which is the analoge to the conditions for kwadratic equations to be solvable.

Theorem 13.79. *Let $\langle X, \langle\rangle\rangle$ be a finite dimensional real (complex) inner product space then if $L \in \operatorname{Hom}(X)$ is a self-adjoint linear operator and $b, c \in \mathbb{R}$ such that $b^2 < 4 \cdot c$ then $L^2 + b \cdot L + c \cdot 1_X$ is a isomorphism.*

Proof. As by 13.40 we have that $L^2 + b \cdot L + c \cdot 1_X$ is linear so that using 13.35 we must only prove that $L^2 + b \cdot L + c \cdot 1_X$ is injective. Now $\forall x \in X$ with $x \neq 0$ we have using the Cauchy Schwarz inequality (see 12.104) that $b \cdot \langle L(x), x\rangle = \langle b \cdot L(x), x\rangle \leqslant \|b \cdot L(x)\| \cdot \|x\| = |b| \cdot \|L(x)\| \cdot \|x\|$ giving by multiplying by -1 that

$$-|b| \cdot \|L(x)\| \cdot \|x\| \leqslant b \cdot \langle L(x), x\rangle \tag{13.52}$$

$$\begin{array}{rcl}\langle (L^2 + b \cdot L + c \cdot 1_X)(x), x\rangle & = & \langle L^2(x) + b \cdot L(x) + c \cdot x, x\rangle \\ & = & \langle L^2(x), x\rangle + b \cdot \langle L(x), x\rangle + c \cdot \langle x, x\rangle \\ & \underset{T \text{isselfadjoint}}{=} & \langle L(x), L(x)\rangle + b \cdot \langle L(x), x\rangle + c \cdot \langle x, x\rangle \\ & = & \|L(x)\|^2 + b \cdot \langle L(x), x\rangle + c \cdot \|x\|^2 \\ & \geqslant_{13.52} & \|L(x)\|^2 - |b| \cdot \|L(x)\| \cdot \|x\| + c \cdot \|x\|^2\end{array}$$

proving

$$\langle (L^2 + b \cdot L + c \cdot 1_X)(x), x\rangle \geqslant \|L(x)\| - |b| \cdot \|L(x)\| \cdot \|x\| + c \cdot \|x\| \tag{13.53}$$

Next

$$\begin{array}{rcl}\left(\|L(x)\| - \frac{|b| \cdot \|x\|}{2}\right)^2 + \left(c - \frac{b^2}{4}\right) \cdot \|x\|^2 & = & \|L(x)\|^2 - |b| \cdot \|L(x)\| \cdot \|x\| + \frac{|b|^2 \cdot \|x^2\|}{4} + c \cdot \|x\|^2 - \frac{b^2 \cdot \|x\|^2}{4} \\ & \underset{b^2 = |b|^2}{=} & \|L(x)\|^2 - |b| \cdot \|L(x)\| \cdot \|x\| + c \cdot \|x\|^2 \\ & \leqslant_{14.157} & \langle (L^2 + b \cdot L + c \cdot 1_X)(x), x\rangle\end{array}$$

Now as $0 \leqslant \left(\|L(x)\| - \frac{|b| \cdot \|x\|}{2}\right)^2$ and $b^2 < 4 \cdot c \Rightarrow 0 < c - \frac{b^2}{4} \underset{0 < \|x\|}{\Rightarrow} 0 < \left(c - \frac{b^2}{4}\right) \cdot \|x\|^2$ we have by the above that

$$0 < \langle (L^2 + b \cdot L + c \cdot 1_X)(x), x\rangle \tag{13.54}$$

From the above we have then that $\forall x \in X \vdash x \neq 0$ $(L^2 + b \cdot L + c \cdot 1_X)(x) \neq 0$ hence $\ker(L^2 + b \cdot L + c \cdot 1_X) = \{0\}$ and thus by 10.211 we have that $L^2 + b \cdot L + c \cdot 1_X$ is injective. □

We are now ready to prove that self adjoint operators have a eigenvector with a eigenvalue.

Theorem 13.80. *Let $\langle X, \langle\rangle\rangle$ be a non trivial real (complex) finite dimensional inner product space and $L \in \operatorname{Hom}(X)$ is a self adjoint operator then L has a eigenvector with a eigenvalue which is real by 13.62.*

Proof. First the theorem is already proved in the complex case (see 13.51) so we have only to consider the real case. Let $n = \dim(X) \in \mathbb{N}$ and choose $x \in X$ with $x \neq 0$. Then as $\{L^i(x)\}_{i \in \{0,\ldots,n\}}$ consists of $n+1$ vectors the family is linear dependent and thus there exists a $\{a_i\}_{i \in 1,\ldots,n} \subseteq \mathbb{R}$ such that $\sum_{i=0}^n a_i \cdot L^i(x) = 0$ with $\{i \in \{0,\ldots,n\} | a_i \neq 0\} \neq \emptyset$ so $N = \max(\{i \in \{1,\ldots,n\}\})$ exists and we have

$$\sum_{i=0}^N a_i \cdot L^i(x) = 0 \text{ with } a_N \neq 0 \text{ where } N \in \{0,\ldots,n\} \tag{13.55}$$

If $N = 0$ then $0 = \sum_{i=0}^0 a_i \cdot L^i(x) = a_0 \cdot L^0(x) = a_0 \cdot x \underset{a_0 \neq x}{\Rightarrow} x = 0$ a contradiction, so we may assume that $N \neq 0$.

$$N \neq 0 \tag{13.56}$$

Hence the polynomial p with real coefficients defined by $p(z) = \sum_{i=0}^N a_i \cdot z^i$ is a non constant polynomial. Using 13.33 there exists a $m, M \in \mathbb{N}_0$, $\{\lambda_i\}_{i \in \{1,\ldots,m\}} \subseteq \mathbb{R}$, $\{b_i\}_{i \in \{1,\ldots,M\}}, \{c_i\}_{i \in \{1,\ldots,M\}} \subseteq \mathbb{R}$ so that $p(z) = (\Pi_{i=1}^m (z - \lambda_i)) \cdot (\prod_{i=1}^M (z^2 + b_i \cdot z + c_i))$ and $\forall i \in \{1,\ldots,M\}$ we have $b_i^2 < 4 \cdot c - i$. Define now $\{p_i\}_{i \in \{1,\ldots,m+M\}}$ by $p_i(z) = \begin{cases} (z - \lambda_i) \text{ if } i \in \{1,\ldots,m\} \\ (z^2 + b_{i-m} \cdot z + c_{i-m}) \text{ if } i \in \{m+1,\ldots,m+M\} \end{cases}$ then $\forall z \in \mathbb{C}$ we have

$$\begin{aligned}
\prod_{i \in \{1,\ldots,m+M\}} p_i(z) &= \prod_{i \in \{1,\ldots,m+M\}} p_i(z) \\
&= \left(\prod_{i \in \{1,\ldots,m\}} p_i(z)\right) \cdot \left(\prod_{i \in \{m+1,\ldots,m+M\}} p_i(z)\right) \\
&= \left(\prod_{i=1}^m (z - \lambda_i)\right) \cdot \left(\prod_{i=m+1}^{m+M} z^2 + b_{i-m} \cdot z + c_{i-m}\right) \\
&= \left(\prod_{i=1}^m (z - \lambda_i)\right) \cdot \left(\prod_{i=1}^M z^2 + b_i \cdot z + c_i\right) \\
&= p(z)
\end{aligned}$$

to summarize we have

$$\forall z \in \mathbb{C} \text{ we have } p(z) = \prod_{i \in \{1,\ldots,m+M\}} p_i(z) \tag{13.57}$$

Using 13.43 we have then that

$$\sum_{i=0}^N a_i \cdot L^i = p(L) = \prod_{i=1}^{m+M} p_i(L) \tag{13.58}$$

Now as by 13.55 we have that $p(L)(x) = \sum_{i=0}^{N} a_i \cdot L^i(x) = 0$ hence as $x \neq 0$ we have that $p(L)$ is not regular. Using 13.58 together with 13.48 proves that there exists a $i \in \{1, \ldots, m + M\}$ such that $p_i(L)$ is not regular. Now if $i \in \{m + 1, \ldots, M + 1\}$ we have $p_i(L) = L^2 + a_i \cdot L + c_i$ which as $b_i^2 < 4 \cdot c_i$ is regular by 13.79, so we must have that $i \in \{1, \ldots, m\}$ hence $p_i(L) = L - \lambda_i \cdot 1_X$ is not regular, so there exists a $y \in X$ with $y \neq 0$ such that $L(y) - \lambda_i \cdot 1_X(y) = 0 \Rightarrow L(y) = \lambda_i \cdot y$. So we have found our eigenvector with eigenvalue λ. □

Definition 13.81. *Let $\langle X, \langle\rangle\rangle$ be a real (complex) inner product vactor space and $U \subseteq X$ then $U^{\perp} = \{x \in x | \forall y \in U$ we have $\langle x, y\rangle = 0\}$. U is called the orthogonal complement of .*

Theorem 13.82. *Let $\langle X, \langle\rangle\rangle$ be a real (complex) inner vactor space then*

1. *If $U \subseteq X$ then $U^{\perp}$ is a sub space of X*
2. *$\{0\}^{\perp} = X$*
3. *$X^{\perp} = \{0\}$*
4. *If $U \subseteq X$ then $U \bigcap U^{\perp} \subseteq \{0\}$*
5. *If $U \subseteq X$ is a subspace then $U \bigcap U^{\perp} = \{0\}$*
6. *If $U, W \subseteq X$ with $U \subseteq W$ then $W^{\perp} \subseteq U^{\perp}$*

Proof.

1. If $x, y \in U^{\perp}$ and $\alpha, \beta \in \mathbb{R}(\mathbb{C})$ then if $u \in U$ we have $\langle \alpha \cdot x + \beta \cdot y, u\rangle = \alpha \cdot \langle x, u\rangle + \beta \cdot \langle y, u\rangle = 0$ so $U^{\perp}$ is a sub space of X
2. As $\forall x \in X$ we have $\langle X, 0\rangle = 0$ we have that $\{0\}^{\perp} = X$
3. First $\forall x \in X$ we have $\langle 0, x\rangle = 0$ we have $0 \in X^{\perp}$, if $y \in X^{\perp}$ then $\forall x \in X$ we have $\langle y, x\rangle = 0$ hence $\langle y, y\rangle = 0$ proving that $y = 0$ so that $X^{\perp} = \{0\}$
4. If $x \in U \bigcap U^{\perp}$ then $x \in U$ and $x \in U^{\perp}$ so that $\langle x, x\rangle = 0$ proving that $x = 0$ so that $U \bigcap U^{\perp} \subseteq \{0\}$
5. If U is a subspace then $0 \in U$, as by (1) $U^{\perp}$ is also a sub space we have that $0 \in U^{\perp}$ so using (4) we have that $0 \in U \bigcap U^{\perp} \subseteq \{0\}$ proving that $U \bigcap U^{\perp} = \{0\}$
6. If $x \in W^{\perp}$ then $\forall y \in U$ we have $y \in W$ so that $\langle x, y\rangle = 0$ hence $x \in U^{\perp}$ □

Theorem 13.83. *Let $\langle X, \langle\rangle\rangle$ be a real(complex) inner product space, U a finite dimensional sub space of X then $X = U \oplus U^{\perp}$*

Proof. We have the following cases to consider for U

***U* is trivial.** then $U = \{0\}$ we have by the previous theorem that $U^{\perp} = X$ so that $U \oplus U^{\perp} = \{0\} \oplus X = X$

***U* is not trivial.** then as U is dinite dimensional there exists a orthonormal basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ for U (see 12.117). Let $x\in X$ then we have that $x=x+y$ where $x=\sum_{i=1}^n \langle x,e_i\rangle\cdot e_i$ and $y=x-\sum_{i=1}^n \langle x,e_i\rangle\cdot e_i$. If now $u\in U$ then we have that $u=\sum_{i=1}^n u\cdot e_i$ and

$$\begin{aligned}\langle y,u\rangle &= \left\langle x-\sum_{i=1}^n \langle x,e_i\rangle\cdot e_i,\sum_{k=1}^n u_k\cdot e_k\right\rangle\\ &= \sum_{k=1}^n \bar{u_k}\cdot\left\langle x-\sum_{i=1}^n \langle x,e_i\rangle\cdot e_i,e_k\right\rangle\\ &= \sum_{k=1}^n \bar{u_k}\cdot\left(\langle x,e_k\rangle-\left\langle\sum_{i=1}^n \langle x,e_i\rangle\cdot e_i,e_k\right\rangle\right)\\ &= \sum_{k=1}^n \bar{u_k}\cdot\left(\langle x,e_k\rangle-\sum_{i=1}^n (\langle x,e_i\rangle\cdot\langle e_i,e_k\rangle)\right)\\ &= \sum_{k=1}^n \bar{u_k}\cdot\left(\langle x,e_k\rangle-\sum_{i=1}^n (\langle x,e_i\rangle\cdot\delta_{i,k})\right)\\ &= \sum_{k=1}^n \bar{u_k}\cdot(\langle x,e_k\rangle-\langle x,e_k\rangle)\\ &= 0\end{aligned}$$

proving that $y\in U^\perp$ and as x is trivially a element of U we have that $x\in U+U^\perp$. So X$=U+U^\perp$ and as by the previous theorem $U\bigcap U^\perp=\{0\}$ we have by 10.366) that $X=U\oplus U^\perp$ □

Theorem 13.84. *Let $\langle X,\langle\rangle\rangle$ be a real (complex) finite dimensional space, $L\in$ Hom(X) a self adjoint operator and $U\subseteq X$ a sub space of X so that $L(U)\subseteq U$ Then*

1. *$L(U^\perp)\subseteq U^\perp$*
2. *$T_{|U}\in\mathrm{Hom}(U)$ is self adjoint in $\langle U,\langle\rangle_{|U\times U}\rangle$*
3. *$T_{|U^\perp}\in\mathrm{Hom}(U^\perp)$ is self adjoint in $\langle U^\perp,\langle\rangle_{U^\perp\times U^\perp}\rangle$*

Proof.

1. If $y\in L(U^\perp)$ then there exists a $x\in U^\perp$ such that $y=L(x)$, if now $u\in U$ we have that $\langle y,u\rangle=\langle L(x),u\rangle\underset{L\,\mathrm{is\,self\,adjoint}}{=}\langle x,L(u)\rangle\underset{x\in U^\perp\wedge L(u)\in U}{=}0$ proving that $y=L(x)\in U^\perp$. Hence $L(U^\perp)\subseteq U^\perp$
2. If $u,v\in U$ then as $T_{|U}(u)=T(u)\in U$ and $T_{|U}(v)=T(v)$so that $\langle T_{|U}(u),v\rangle_{U\times U}=\langle T_{|U}(u),v\rangle=\langle T(u),v\rangle=\langle u,T(v)\rangle=\langle u,T_{|U}(v)\rangle=\langle u,T_{|U}(n)\rangle_{|U\times U}$ proving that $T_{|U}$ is self adjoint in $\langle U,\langle\rangle_{U\times U}\rangle$

3. This follows from (2) by replacing U by $U^{\perp}$ and the fact that $U^{\perp}$ is also a subspace in X (see 13.82). □

Theorem 13.85. (Real Spectral Theorem) *Let $\langle X, \langle\rangle\rangle$ be a real finite dimensional inner product space with $\dim(X) = n \in \mathbb{N}$ and $L \in \mathrm{Hom}(X)$ then the following are equivalent*

1. *L is self-adjoint*
2. *There exists a orthonormal basis $\{e\}_{i\in\{1,\ldots,n\}}$ for X of eigenvectors of L.*
3. *There exists a basis for X that $\mathcal{M}(L, \{e_i\}_{i\in\{1,\ldots,n\}}, \{e_i\}_{i\in\{1,\ldots,n\}})$ is diagonal.*

Proof.

$1 \Rightarrow 2$. We prove first by induction on the dimension of X that X has a orthonormal basis of egienvectors of T. So let $\mathcal{F} = \{n \in \mathbb{N} |$if $\dim(X) = n$ and $L \in \mathrm{Hom}(X)$ is self-adjoint then there exist a orthonormal basis consisting of eigenvectors of $T\}$ then we have:

$1 \in \mathcal{F}$. If $\dim(X) = 1$ then X has a basis $\{f_i\}_{i\in\{1,\ldots,1\}}$ consisting of one vector f_1, take then $e_1 = \frac{f_1}{\|f_1\|}$ then $\forall x \in X$ there exists a $\alpha_1 \in \mathbb{R}$ such that $x = \alpha_1 \cdot f_1 = \frac{\alpha_1 \cdot \|f_1\|}{\|f_1\|} \cdot f_1 = \alpha_1 \cdot e_1$ proving that $\{e_i\}_{i\in\{1,.,1\}}$ is a orthonormal basis of X. Hence as $L(e_1) \in X$ there exists a $\lambda \in \mathbb{R}$ such that $L(e_1) = \lambda \cdot e_1$ proving that e_1 is a eigenvector of X. So $1 \in \mathcal{F}$

$n \in \mathcal{F} \Rightarrow n + 1 \in \mathcal{F}$. Let $\dim(X) = n + 1$ and $L \in \mathrm{Hom}(X)$ is a self-adjoint operator. Using 13.80 there exists a eigen vector $x \in X$ such that $L(x) = \lambda \cdot x$, as by definition of eigenvectors $x \neq 0$ we can define $e_{n+1} = \frac{x}{\|x\|}$ so that $L(e_{n+1}) = \frac{1}{\|x\|} \cdot L(x) = \frac{\lambda}{\|x\|} \cdot x = \lambda \cdot e_{n+1}$. So we have

$$L \text{ has a eigenvector } e_{n+1} \text{ with } \|e_{n+1}\| = 1 \tag{13.59}$$

Define now $U = \mathcal{S}\{e_{n+1}\} = \{\lambda \cdot e_{n+1} | \lambda \in \mathbb{R}\}$ which is a vector space with $\dim(U) = 1$. Take the vector space $U^{\perp}$ (see 13.82) then by 13.83 we have that $X = U \oplus U^{\perp}$ hence using 10.370 we get $n + 1 = \dim(X) = \dim(U) + \dim(U^{\perp}) = 1 + \dim(U)$. So $\dim(U^{\perp}) = n$. As by the previous theorem (see 13.84) that $L_{|U^{\perp}}$ is self-adjoint, hence using the fact that $n \in \mathcal{F}$ there exists a orthonormal basis $\{e_i\}_{i\in\{1,\ldots,n\}}$ for $U^{\perp}$ of eigen vectors of $L_{|U^{\perp}}$ with eigenvalues $\{\lambda_i\}_{i\in\{1,\ldots,n\}}$. Take now $\{e_i\}_{i\in\{1,\ldots,n+1\}}$ then for $i, j \in \{1, \ldots, n + 1\}$ with $i \neq j$ then we have the following cases:

$i, j \in \{1, \ldots, n\}$. as $\{e_i\}_{i\in\{1,\ldots,n\}}$ is orthonormal we have that $\langle e_i, e_j\rangle = 0$

$i = n + 1, j \in \{1, \ldots, n\}$. as $e_j \in U^{\perp}$ and $e_i \in U$ we have that $\langle e_i, e_j\rangle = 0$

$i \in \{1, \ldots, n\}, j = n + 1$. as $e_i \in U^{\perp}$ and $e_j \in U$ we have tha t $\langle e_i, e_j\rangle = 0$

The above proves that $\{e_i\}_{i\in\{1,\ldots,n+1\}}$ is orthonormal and by 12.118 that it is also a basi of C. Finally $\forall i\in\{1,\ldots,n\}$ we have that $L(e_i) \underset{e_i\in U^\perp}{=} L_{|U^\perp}(e_i)=\lambda_i\cdot e_i$ proving together with 13.59 that $\{e_i\}_{i\in\{1,\ldots,n+1\}}$ is a orthonormal basis for X of eigenvectors of L. This proves that $n+1\in\mathcal{F}$.

2 ⇔ 3. This follows from 13.76.

2 ⇒ 1. Let $\{e_i\}_{i\in\{1,\ldots,n\}}$ be a basis $\{e\}_{i\in\{1,\ldots,n\}}$ for X of eigen vectors for L with eigenvalues $\{\lambda_i\}_{i\in\{1,\ldots,n\}}$. Then if $x\in X$ we have that $x=\sum_{i=1}^n x_i\cdot e_i$, $y=\sum_{i=1}^n y_i\cdot e_i$

$$\begin{aligned}
\langle L(x),y\rangle &= \left\langle L\left(\sum_{i=1}^n x_i\cdot e_i\right),\sum_{j=1}^n y_j\cdot e_j\right\rangle\\
&= \left\langle \sum_{i=1}^n x_i\cdot L(e_i),\sum_{j=1}^n y_j\cdot e_j\right\rangle\\
&= \sum_{i=1}^n x_i\left\langle \lambda_i\cdot e_i,\sum_{j=1}^n y_j\cdot e_j\right\rangle\\
&= \sum_{i=1}^n x_i\cdot\lambda_i\left\langle e_i,\sum_{j=1}^n y_j\cdot e_j\right\rangle\\
&= \sum_{i=1}^n \left(x_i\cdot\lambda_i\cdot\sum_{j=1}^n y_j\cdot\langle e_i,e_j\rangle\right)\\
&= \sum_{i=1}^n \lambda_i\cdot x_i\cdot y_i\\
\langle x,L(y)\rangle &= \left\langle \sum_{i=1}^n x_i\cdot e_i, L\left(\sum_{j=1}^n y_j\cdot e_j\right)\right\rangle\\
&= \sum_{i=1}^n x_i\cdot\left\langle e_i,\sum_{j=1}^n y_j\cdot L(e_j)\right\rangle\\
&= \sum_{i=1}^n x_i\cdot\left(\sum_{j=1}^n y_j\cdot\lambda_j\cdot\langle e_i,e_j\rangle\right)\\
&= \sum_{i=1}^n \lambda_i\cdot x_i\cdot y_i\\
&= \langle L(x),y\rangle
\end{aligned}$$

proving that L is self-adjoint □

Corollary 13.86. *Let $\langle X,\langle\rangle\rangle$ be a real finite dimensional inner product space with $\dim(X)=n\in\mathbb{N}$, $\{e_i\}_{i\in\{1,\ldots,n\}}$ a orthonormal base, $L\in\mathrm{Hom}(X)$ a adjoint operator then there exists a unitary operator $U\in\mathrm{Hom}(X)$ such that $\mathcal{M}(U^*\circ L\circ U$, $\{e_i\}_{i\in\{1,\ldots,n\}})$ is diagonal.*

Proof. Using the Real spectral theorem 13.85 there exists a basis $\{f_i\}_{i\in\{1,\ldots,n\}}$ such that $\mathcal{M}(L,\{f_i\}_{i\in\{1,\ldots,n\}})$ is diagonal. Take now the unitary mapping $U\in\operatorname{Hom}(X)$ defined by $\forall i\in\{1,\ldots,n\}$ $U(e_i)=f_i$ (see 13.70) then we have $\forall i,j\in\{1,\ldots,n\}$ that

$$\begin{aligned}\mathcal{M}(U^*\circ L\circ U,\{e_i\}_{i\in\{1,\ldots,n\}})_{i,j} &\underset{12.115}{=} \langle (U^*\circ L\circ U)(e_j),e_i\rangle\\ &= \langle U^*(L(U(e_j))),e_i\rangle\\ &= \langle U^*(L(f_j)),e_i\rangle\\ &= \langle L(f_j),(U^*)^*(e_i)\rangle\\ &= \langle L(f_j),U(e_i)\rangle\\ &= \langle L(f_j),f_i\rangle\\ &\underset{12.115}{=} \mathcal{M}(L,\{f_i\}_{i\in\{1,\ldots,n\}})_{i,j}\\ &= \mathcal{M}(L,\{f_i\}_{i\in\{1,\ldots,n\}})_{i,i}\cdot\delta_{i,j}\end{aligned}$$

□

Theorem 13.87. *Let $n\in\mathbb{N}$ then we have we have*

1. *If $M\in\mathcal{M}(n\times n,\mathbb{R})$ such that M is symmetric then there exists a unitary matrix U such that $U^T\cdot M\cdot U\underset{U\text{ is unitary}}{=}U^{-1}\cdot M\cdot U$ is diagonal*
2. *If $M\in\mathcal{M}(n\times n,\mathbb{C})$ such that M is Hermitian then there exists a unitary matrix U such that $U^T\cdot M\cdot U\underset{U\text{ is unitary}}{=}U^{-1}\cdot M\cdot U$ is diagonal.*

Proof.

1. Take the inner product vector space $\langle\mathbb{R}^n,\langle\rangle\rangle$ with the orthonormal basis $\{\varepsilon_i\}_{i\in\{1,\ldots,n\}}$ (see 12.113) and define $L:\mathbb{R}^n\to\mathbb{R}^n$ by $L(\varepsilon_i)=\sum_{k=1}^n M_{k,i}\cdot\varepsilon_k$ so that $\mathcal{M}(L,\{\varepsilon_i\}_{i\in\{1,\ldots,n\}})=M$. As M is symmetric we have in the real case that M is Hermitian and thus by 13.59 that L is self-adjoint. Using the Real Spectral theorem (see 13.85) there exists a orthonormal basis $\{f_i\}_{i\in\{1,\ldots,n\}}$ such that $\mathcal{M}(L,\{f_i\}_{i\in\{1,\ldots,n\}})$ is diagonal. Using 13.70 we have then that $\mathcal{M}(L,\{f_i\}_{i\in\{1,\ldots,n\}})=\mathcal{M}(U,\{\varepsilon_i\}_{i\in\{1,\ldots,n\}})^{-1}\cdot\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}})\cdot\mathcal{M}(U,\{\varepsilon_i\}_{i\in\{1,\ldots,n\}})=U^{-1}\cdot M\cdot U$ where $U=\mathcal{M}(U,\{\varepsilon_i\}_{i\in\{1,\ldots,n\}})$ is unitary (see 13.70 and the fact that $\{\varepsilon_i\}_{i\in\{1,\ldots,n\}}$ and $\{f_i\}_{i\in\{1,\ldots,n\}}$ are orthonormal).
2. Take the inner product vector space $\langle\mathbb{C}^n,\langle\rangle\rangle$ with the orthonormal basis $\{\varepsilon_i\}_{i\in\{1,\ldots,n\}}$ (see 12.113) and define $L:\mathbb{C}^n\to\mathbb{C}^n$ defined by $L(\varepsilon_i)=\sum_{k=1}^n M_{k,i}\cdot\varepsilon_k$ so that $\mathcal{M}(L,\{\varepsilon_i\}_{i\in\{1,\ldots,n\}})=M$. As M is Hermitian we have by 13.59 that L is self-adjoint and thus rormal (see 13.72). Using the Complex Spectral theorem (see 13.85) there exists a orthonormal basis $\{f_i\}_{i\in\{1,\ldots,n\}}$ such that $\mathcal{M}(L,\{f_i\}_{i\in\{1,\ldots,n\}})$ is diagonal. Using 13.70 we have then that $\mathcal{M}(L,\{f_i\}_{i\in\{1,\ldots,n\}})=\mathcal{M}(U,\{\varepsilon_i\}_{i\in\{1,\ldots,n\}})^{-1}\cdot\mathcal{M}(L,\{e_i\}_{i\in\{1,\ldots,n\}})\cdot\mathcal{M}(U,\{\varepsilon_i\}_{i\in\{1,\ldots,n\}})=U^{-1}\cdot M\cdot U$ where $U=\mathcal{M}(U,\{\varepsilon_i\}_{i\in\{1,\ldots,n\}})$ is unitary (see 13.70 and the fact that $\{\varepsilon_i\}_{i\in\{1,\ldots,n\}}$ and $\{f_i\}_{i\in\{1,\ldots,n\}}$ are orthonormal). □

Chapter 14

Differentiability in normed vector spaces

We assume in this chapter that, unless stated otherwise, all the normed and Banach spaces are using the same field $\mathbb{K}$ where $\mathbb{K}$ is either $\mathbb{C}$ or $\mathbb{R}$.

14.1 Differentiability

14.1.1 Differential

A function is differentiable at a point if we can aproximate the function with a continuous linear function in a small enough neighborhood around the point. To ensure that we don't evaluate the function outside its domain we define given a set U the set U_x.

Definition 14.1. *Let $\langle X, \|\|\rangle$ be a normed vector space and let $U \subseteq X, x \in U$ then $U_x = \{h \in X | x + h \in U\}$*

Theorem 14.2. *Let $\langle X, \|\|\rangle$ be a normed vector space, $U \subseteq X$, $x \in U$ then $U_x = (-x) + U = U - x$*

Proof. First

$$\begin{aligned} h \in U_x &\Leftrightarrow x + h \in U \\ &\Leftrightarrow \exists y \in U \text{ with } x + h = y \\ &\Leftrightarrow \exists y \in U \text{ with } h = (-x) + y \\ &\Leftrightarrow h \in (-x) + U \end{aligned}$$

Second

$$\begin{aligned} h \in (-x) + U &\Leftrightarrow \exists y \in U \text{ with } h = (-x) + y \\ &\Leftrightarrow \exists y \in U \text{ with } h = y - x \\ &\Leftrightarrow h \in U - x \end{aligned}$$

□

Theorem 14.3. *Let $\langle X, \|\|\rangle$ be a normed vector space and let $U \subseteq X$ open, $x \in U$ then U_x is open.*

Proof. This follows directly from the above and 12.76 $\square$

Let's now define the differentiability of a function.

Definition 14.4. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be real normed spaces, $U \subseteq X$ a open set then a function $f: U \rightarrow Y$ is differentiable at $x \in U$ if and only if there exists a continuous linear function $L: X \rightarrow Y$ (or $L \in L(X, Y)$) such that $\forall \varepsilon > 0$ there exists a $\delta > 0$ such that for all $h \in U_x$ with $\|h\|_x < \delta$ we have $\|f(x+h) - f(x) - L(h)\|_Y \leqslant \varepsilon \cdot \|h\|_X$*

Note that $h \in U_x$ ensures that $x + h \in U = \operatorname{dom}(f)$ so our definition is well defined. Next we will prove that the linear aproximation of f is unique.

Theorem 14.5. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be real normed space, $U \subseteq X$ a open set, $x \in U$ and $f: U \rightarrow Y$ differentiable at x then the L used in the above definition is unique*

Proof. Let $h \in X$ then we have the following cases to consider:

$h = 0$. then $\|L_1(h) - L_2(h)\|_Y = \|L_1(0) - L_2(0)\|_Y = \|0 - 0\|_{|Y} = 0$

$h \neq 0$. Let $\varepsilon > 0$ then for $i \in \{1, .2\}$ there exists $\delta_i > 0$ such that for $h \in U_x$ and $\|h\|_X < \delta_i$ we have $\|f(x+h) - f(x) - L_i(h)) \leqslant \frac{\varepsilon}{2} \cdot \|h\|_X$. Further as U_x is open there exists a $\delta_3 > 0$ such that $B_{\|\|_X}(x, \delta_3) \subseteq U_x$. Take $\delta = \min(\delta_1, \delta_2, \delta_3)$ and define $h' = \frac{\delta}{2 \cdot \|h\|_X} \cdot h$ then $\|h'\|_X = \frac{\delta}{2 \cdot \|h\|_X} \cdot \|h\|_X = \frac{\delta}{2} < \delta_1, \delta_2, \delta_3$ so that $h' \in U_x$ and

$$\begin{aligned} \|L_1(h') - L_2(h')\|_Y &= \|(f(x+h') - f(x') - L_2(h')) - (f(x+h') - f(x) - \\ &\quad L_1(h'))\|_Y \\ &\leqslant \frac{\varepsilon}{2} \cdot \|h'\|_X + \frac{\varepsilon}{2} \cdot \|h'\|_X = \varepsilon \cdot \|h'\|_X \end{aligned} \tag{14.1}$$

hence

$$\begin{aligned} \frac{\delta}{2 \cdot \|h\|_X} \cdot \|L_1(h) - L_2(h)\|_Y &\underset{L_1 L_2 \text{are linear}}{=} \left\| L_2\left(\frac{\delta}{2 \cdot \|h\|_X} \cdot h\right) - L_1\left(\frac{\delta}{2 \cdot \|h\|_X} \cdot h\right) \right\|_Y \\ &= \|L_2(h') - L_1(h')\|_Y \\ &\leqslant_{14.1} \varepsilon \cdot \|h'\|_X \end{aligned}$$

As ε was chosen arbitrary we have by 9.56 that $0 \leqslant \frac{\delta}{2 \cdot \|h\|_X} \cdot \|L_1(h) - L_2(h)\|_Y \leqslant 0$ proving that $\|L_1(h) - L_2(h)\|_Y = 0$.

So we have for $h \in X$ that $\|L_1(h) - L_2(h)\|_Y = 0 \Rightarrow L_1(h) - L_2(h) = 0 \Rightarrow L_1(h) = L_2(h)$ which proves that $L_1 = L_2$. $\square$

The unique continuous linear aproximation of a function at a point is called the differential of the function at this point.

Definition 14.6. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be real normed spaces, $U \subseteq X$ a open set, $x \in U$ and $f: U \rightarrow Y$ differentiable at x then the unique (by the previous theorem) $L \in L(X, Y)$ such that $\forall \varepsilon > 0$ there exists a $\delta > 0$ so that if $h \in U_x$ and $\|h\|_X < \delta$ then $\|f(x+h) - f(x) - L(h)\|_Y \leqslant \varepsilon \|h\|_X$ is called the **differential of f at x** and is noted as $Df(x)$. In other words f is differentiable at x with differential $Df(x)$ if and only there exists a $Df(x) \in L(X, Y)$ (which is unique) such that $\forall \varepsilon > 0$ there exists a $\delta > 0$ such that if $h \in U_x$ and $\|h\|_X < \delta$ then $\|f(x+h) - f(x) - Df(x)(h)\|_Y \leqslant \varepsilon \cdot \|h\|_X$*

Let's now look at some alternative definitions for differentiability at a point.

Definition 14.7. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces, $U \subseteq X$ a open set and $x \in U$ then a ε-mapping at x is a function $\varepsilon: U_x \rightarrow Y$ which is continuous at $0 \in U_x$ (using the subspace topology on U_x) and for which $\varepsilon(0) = 0$*

We prove now the following equivalent alternative definitions of differentiability

Theorem 14.8. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces, $U \subseteq X$ a open set, $x \in U$ and $f: U \rightarrow Y$ a function then the following are equivalent*

1. *f is differentiable at x with differential $Df(x)$*
2. *There exists a linear function $Df(x): X \rightarrow Y$ and a ε-mapping $\varepsilon: U_x \rightarrow Y$ such that $\forall h \in U_x$ we have $f(x+h) - f(x) - Df(x)(h) = \varepsilon(h) \cdot \|h\|_X$*
3. *There exists a linear function $Df(x): X \rightarrow Y$ such that $\forall \varepsilon > 0$ there exists a $\delta > 0$ such that $\forall h \in U_x$ with $0 < \|h\| < \varepsilon$ we have $\frac{\|f(x+h) - f(x) - Df(x)(h)\|_Y}{\|h\|_X} < \varepsilon$*
4. *There exists a linear function $Df(x): X \rightarrow Y$ such that $\forall \{h_i\}_{i \in \mathbb{N}}$ with $\lim_{h \rightarrow 0} h_i = 0$ and $\forall i \in \mathbb{N}$ $\|h_i\|_X > 0$ and $h_i \in U_x$ we have that $\lim_{n \rightarrow \infty} \frac{\|f(x+h_i) - f(x) - Df(x)(h_i)\|_Y}{\|h_i\|_X} = 0$*

Proof.

$1 \Rightarrow 2$. Define $\varepsilon: U_x \rightarrow Y$ by $\varepsilon(h) = \begin{cases} 0 \text{ if } h = 0 \\ \frac{f(x+h) - f(x) - Df(x)(h)}{\|h\|_X} \text{ if } h \neq 0 \end{cases}$ then we have that $\varepsilon(h) \cdot \|h\|_X = \begin{cases} 0 \text{ if } h = 0 \\ f(x+h) - f(x) - Df(x)(h) \text{ if } h \neq 0 \end{cases} = f(x+h) - f(x) = Df(x) \cdot \|h\|_x$. Further take $\zeta > 0$ then by definition there exists a $\delta > 0$ such that for $h \in U_x$ with $\|h\|_X$ we have that $\|f(x+h) - f(x) - Df(x)(h)\|_Y \leqslant \frac{\zeta}{2} \cdot \|h\|_X < \zeta \cdot \|h\|_X$ so that

$$\begin{aligned} \|\varepsilon(h) - \varepsilon(0)\|_Y &= \|\varepsilon(h)\|_Y \\ &= \begin{cases} \|0\|_Y \text{ if } h < \zeta \\ \frac{\|f(x+h) - f(x) - Df(x)(h)\|_Y}{\|h\|_X} < \zeta \end{cases} \\ &\leqslant \zeta \end{aligned}$$

which using 12.75 and 12.151 proves that ε is continuous at 0 in the subspace topology on U_x.

2 ⇒ 3. Let $\varepsilon: U_x \to Y$ be the ε-mapping such that $f(x+h) - f(x) - Df(x)(x) = \varepsilon(h) \cdot \|h\|_X \ \forall h \in U_x$. Then as ε is continuous at 0 and $\varepsilon(0) = 0$ there exists given a $\zeta > 0$ a $\delta > 0$ such that if $h \in U_x$ with $\|h\|_X < \delta$ we have $\|\varepsilon(h)\|_Y = \|\varepsilon(h) - 0\|_Y = \|\varepsilon(h) - \varepsilon(0)\|_Y < \zeta$ so that if additional $0 < \|h\|_X$ we have $\frac{\|f(x+h) - f(x) - Df(x)(h)\|_Y}{\|h\|_X} = \frac{\|\varepsilon(h)\| \cdot \|h\|_X}{\|h\|_X} = \|\varepsilon(h)\|_X < \zeta$

3 ⇒ 1. Let $\varepsilon > 0$, find a $\delta > 0$ such that for $h \in U_x$ with $0 < \|h\|_X < \delta$ we have $\frac{\|f(x+h) - f(x) - Df(x)(h)\|_Y}{\|h\|_Y} < \varepsilon$. Now $\forall h \in U_x$ with $\|h\|_X < \delta$ we have two cases to consider

$\|\boldsymbol{h}\|_{\boldsymbol{X}} = \mathbf{0}$. then $h = 0$ so that $\|f(x+h) - f(x) - Df(x)(h)\|_Y = \|f(x) - f(x) - Df(x)(0)\|_Y = \|0 - 0\|_Y = 0 \leqslant 0 = \varepsilon \cdot 0 = \varepsilon \cdot \|h\|_X$

$\|\boldsymbol{h}\|_{\boldsymbol{X}} \neq \mathbf{0}$. then $\frac{\|f(x+h) - f(x) - Df(x)(h)\|_Y}{\|h\|_X} < \varepsilon \Rightarrow \|f(x+h) - f(x) - Df(x)(x)\| \leqslant \varepsilon \cdot \|h\|_X$

so in all cases we have $\|f(x+h) - f(x) - Df(x)(h)\|_Y \leqslant \varepsilon \cdot \|h\|_X$ proving differentiability at x.

3 ⇒ 4. Let $\{h_i\}_{i \in \mathbb{N}}$ be a sequence of strict positive real numbers with limit 0. Take now $\varepsilon > 0$ then by (3) there exists a δ_ε such that $\forall h \in U_x$ with $0 < \|h\|_X < \delta_\varepsilon$ we have $\frac{\|f(x+h) - f(x) - Df(x)(h)\|_Y}{\|h\|_X} < \varepsilon$. As $\lim_{n \to \infty} h_n = 0$ there exists a $N \in \mathbb{N}$ such that $\forall n \geqslant N$ we have $\|h_n\| < \delta_\varepsilon$, then as we have assumed that $\forall n \in \mathbb{N} \ \ 0 < h_n \in U_x$, we have if $n \geqslant N$ that $\frac{\|f(x+h_n) - f(x) - Df(x)(h_n)\|_Y}{\|h_n\|_X} < \varepsilon$ proving that

$$\lim_{n \to \infty} \frac{\|f(x+h_n) - f(x) - Df(x)(h)\|_Y}{\|h_n\|_X} = 0$$

4 ⇒ 3. Assume that (3) is false then there exists a $\varepsilon_0 > 0$ such that $\forall \delta > 0$ there exists a $h \in U_x$ with $0 < \|h\|_X < \delta$ so that $\frac{\|f(x+h) - f(x) - Df(x)(h)\|_Y}{\|h\|_X} > \varepsilon_0$. So for every $n \in \mathbb{N}$ there exists a $h_n \in U_x$ with $0 < \|h_n\|_X < \frac{1}{n}$ so that $\frac{\|f(x+h_n) - f(x) - Df(x)(h_n)\|_Y}{\|h_n\|_X} > \varepsilon_0$. Now if $\delta > 0$ there exists by the Archimedean property (see 9.55) a $N_\delta \in \mathbb{N}$ such that $\frac{1}{N_\delta} < \delta$ so if $n \geqslant N_\delta$ we have $\|h_n - 0\|_X = \|h_n\|_X < \frac{1}{n} \leqslant \frac{1}{N_\delta} < \delta$ proving that $\lim_{n \to \infty} h_n = 0$.Now by the hypothese (4) we must have that $\lim_{n \to \infty} \frac{\|f(x+h_n) - f(x) - Df(x)(h_n)\|_Y}{\|h_n\|_X} = 0$ and thus there exists a $N \in \mathbb{N}$ such that $\frac{\|f(x+h_N) - f(x) - Df(x)(h_N)\|_Y}{\|h_N\|} < \varepsilon_0$ contradicting $\frac{\|f(x+h_N) - f(x) - Df(x)(h_N)\|_Y}{\|h_N\|} > \varepsilon_0$. As the assumption gives a contradiction we must have that (3) is true. □

Remark 14.9. If $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ are finite dimensional spaces then by 12.296 we have that every linear map between X and Y is continuous, so in the the previous theorem and the definition of differentiability we can leave out the requirement for the continuity of $Df(x)$ in the finite dimensional case.

The following theorem is a consequence of the fact that the differential is a continuous map.

Theorem 14.10. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces, $U \subseteq X$ a open set, $x \in U$and $f: U \rightarrow Y$ a function that is differentiable at x then f is continuous at x.*

Proof. Let W be a open subset of Y with $f(x) \in W$ then we have

$$\exists \delta_1 > 0 \,\text{such that}\, f(x) \subseteq B_{\|\|_Y}(f(x), \delta_1) \subseteq W \tag{14.2}$$

From the alternative definition of differentiability at x (see 14.8) there exists a ε-mapping $\varepsilon: U_x \rightarrow Y$ such that $\forall h \in U_x$ we have $f(x+h) - f(x) - Df(x)(h) = \varepsilon(h)$. As ε is continuous at 0 and $\varepsilon(0) = 0$ we have

$$\exists \delta_2 > 0 \text{ such that if } h \in U_x \text{ and } \|h\|_X < \delta_2 \text{ then } \|f(x+h) - f(x) - Df(x)(h)\|_Y < \frac{\delta_1}{2} \tag{14.3}$$

Using the continuity of the linear map $Df(x)$ and $Df(x)(0) = 0$ we have also

$$\exists \delta_3 > 0 \,\text{such that if}\, \|h\|_X < \delta_3 \,\text{then}\, \|Df(x)(h)\|_Y < \frac{\delta_1}{2} \tag{14.4}$$

As U is open and $x \in U$ we have

$$\exists \delta_4 > 0 \,\text{such that}\, x \in B_{\|\|_X}(x, \delta_4) \subseteq U \tag{14.5}$$

Define now $\delta = \min(\delta_1, \delta_2, \delta_3, \delta_4)$. If $y \in B_{\|\|_X}(x, \delta) \subseteq B_{\|\|_X}(x, \delta_4)$ then by the above $y \in U$ so that $(y - x) + x = y \in U$ or $y - x \in U_x$ and as $\|y - x\|_X < \delta < \delta_2$ we have by 14.3 that $\|f(x + (y - x)) - f(x) - Df(x)(y - x)\|_Y < \frac{\delta_1}{2}$ giving

$$\forall y \in B_{\|\|_X}(x, \delta) \,\text{we have}\, \|f(y) - f(x) - Df(x)(y - x)\|_Y < \frac{\delta_1}{2} \tag{14.6}$$

Further if $y \in B_{\|\|_X}(x, \delta)$ we have also that $\|y - x\|_X < \delta_3$ so that by 14.4 we have

$$\forall y \in B_{\|\|_X}(x, y) \,\text{we have}\, \|Df(x)(y - x)\|_Y < \frac{\delta_1}{2} \tag{14.7}$$

Finally if $y \in B_{\|\|_X}(x, \delta)$ then we have

$$\begin{aligned}
\|f(y) - f(x)\|_Y &= \|f(y) - f(x) - Df(x)(y - x) + Df(x)(y - x)\|_Y \\
&\leqslant \|f(y) - f(x) - Df(x)(y - x)\|_Y + \|Df(x)(y - x)\|_Y \\
&<_{14.6.14.7} \frac{\delta_1}{2} + \frac{\delta_1}{2} = \delta
\end{aligned}$$

which proves that f is continuous at x. □

The next theorem shows that differentiation is independent of the chosen norm (as long as the norms are equivalent).

Theorem 14.11. *Let X,Y be vector spaces over $\mathbb{K}$ with equivalent norms $\|\|_{X_1}$, $\|\|_{X_2}$ on X, equivalent norms $\|\|_{Y_1}$, $\|\|_{Y_2}$ on Y, $U\subseteq X$ a open set (as norms are equivalent the topology on X,Y is the same for the two norms), $f:U\rightarrow Y$ a function then if f is differentiable at $x\in U$ with differential $Df(x)$ (using norms $\|\|_{X_1}, \|\|_{Y_1}$) we have that f is differentiable at $x\in U$ with differential $Df(x)$ (using norms $\|\|_{X_2}$, $\|\|_{Y_2}$)*

Proof. From the equivalence of the normes we have by 12.82 that there exists α_X, $\beta_X, \alpha_Y, \beta_Y>0$ such that $\forall x\in X$ and $\forall y\in Y$ we have

$$\alpha_X\cdot\|x\|_{X_2}\leqslant\|x\|_{X_1}\leqslant\beta_X\cdot\|x\|_{X_2} \text{ and } \forall y\in Y\vDash\alpha_Y\cdot\|y\|_{Y_1}\leqslant\|y\|_{Y_2}\leqslant\beta_Y\cdot\|y\|_{Y_1} \tag{14.8}$$

Assume that $f:U\rightarrow Y$ is differentiable at x with differential $Df(x)$ using the norms $\|\|_{X_1}$,$\|\|_{Y_1}$ then for $\varepsilon>0$ there exists a δ' such that

$$\|f(x+h)-f(x)-Df(x)(h)\|_{Y_1}\leqslant\frac{\varepsilon}{\beta_Y\cdot\beta_X}\cdot\|h\|_{X_1} \tag{14.9}$$

Take $\delta=\frac{\delta'}{\beta_X}$ then $\forall h\in U_x$ with $\|h\|_{X_2}<\delta$ we have $\beta_X\cdot\|h\|_{X_2}<\delta' \underset{14.8}{\Rightarrow} \|x\|_{X_1}<\delta'$ so using 14.9 we have that $\|f(x+h)-f(x)-Df(x)(h)\|_{Y_1}\leqslant\frac{\varepsilon}{\beta_Y\cdot\beta_X}\cdot\|h\|_{X_1}$ so that $\beta_Y\cdot\|f(x+h)-f(x)-Df(x)(h)\|_{Y_1}\leqslant\frac{\varepsilon}{\beta_X}\cdot\|h\|_{X_1}$ or using 14.8 we have

$$\|f(x+h)-f(x)-Df(x)(h)\|_{Y_2}\leqslant\varepsilon\cdot\|h\|_{X_2}$$

which together with the fact that $Df(x)$ is also a continuous function between X an Y (as the topologies are the same) proves that f has also a differential $Df(x)$ at x using the norms $\|\|_{X_2}$ and $\|\|_{Y_2}$. □

We now prove that differentiability of $f:U\rightarrow V$ at x is essential a local property and is not dependend on U but just on the value of f near x.

Theorem 14.12. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces, $U\subseteq X$ a open set, $x\in U$ and $f:U\rightarrow Y$ then*

1. *If $f:U\rightarrow Y$ is differentiable at x then $\forall W$ open with $x\in W\subseteq U$ we have that $f_{|W}:W\rightarrow Z$ is differentiable at x.*
2. *If there exists a $W\subseteq X$ open with $x\in W\subseteq U$ such that $f_{|W}:W\rightarrow Z$ is differentiable at x then $f:W\rightarrow V$ is differentiable at x.*

Also $Df(x)=D(f_{|V})(x)$ in (1) and (2).

Proof.

1. As f is differentiable we have given $\varepsilon>0$ a $\delta>0$ so that $\forall h\in U_x$ with $\|h\|_X<\delta$ we have $\|f(x+h)-f(x)-Df(x)(h)\|_Y\leqslant\varepsilon\cdot\|h\|_X$. Now if $h\in W_x\Rightarrow x+h\in W\subseteq U\Rightarrow h\in U_x\Rightarrow W_x\subseteq U_x$. So if $h\in W_x$ with $\|h\|_X<\delta$ we have $\|f_{|W}(x+h)-f_{|W}(x)-Df(x)(h)\|_Y \underset{x+h,x\in W}{=} \|f(x+h)-f(x)-Df(x)(h)\|_Y\leqslant\varepsilon\cdot\|h\|_X$ proving that $f_{|W}:W\rightarrow Z$ is differentiable with differential $Df(x)$ and thus $D(f_{|W})(x)=Df(x)$.

2. First as $x \in W$ open there exists a $\delta_1 > 0$ such that $x \in B_{\|\|_X}(x, \delta_1) \subseteq W$, so if $\|h\|_X \leqslant \delta_1$ then $\|x + h - x\|_X = \|h\|_X < \delta_1 \Rightarrow x + h \in B_{\|\|_X}(x, \delta_1) \subseteq W$ or in other words

$$\exists \delta_1 > 0 \text{ such that if } \|h\|_X \leqslant \delta_1 \Rightarrow h \in W_x \tag{14.10}$$

Take now $\varepsilon > 0$, then by differentiability of $f_{|W}$ there exists a $\delta_2 > 0$ such that if $h \in W_x$ and $\|h\|_X < \delta_2$ then $\|f_{|W}(x+h) - f_{|W}(x) - D(f_{|W})(x)(h)\|_Y \leqslant \varepsilon \cdot \|h\|_X \Rightarrow \|f(x+h) - f(x) - D(f_{|W})(x)(h)\|_{|Y} \leqslant \varepsilon \cdot \|h\|_X$ or in other words

$$\text{if } h \in W_x \text{ and } \|h\|_X < \delta_2 \text{ then } \|f(x+h) - f(x) - D(f_{|W})(x)(h)\|_Y \leqslant \varepsilon \cdot \|h\|_X \tag{14.11}$$

Take now $\delta = \min(\delta_1, \delta_2)$ then if $h \in U_x$ and $\|h\|_X < \delta$ then $\|h\|_X < \delta_1 \underset{14.10}{\Rightarrow} h \in W_x$ and as $\|h\|_X < \delta_2$ we have by 14.11 that $\|f(x+h) - f(x) - D(f_{|W})(x)(h)\|_Y \leqslant \varepsilon \cdot \|h\|_X$ proving that f is differentiable at x with differential $D(f_{|W})(x)$ thus $Df(x) = D(f_W)(x)$. □

Using the alternative definitions of differentiability (see 14.8) and the above theorem we have then the following criteria for differentiability.

Corollary 14.13. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be normed spaces, $U \subseteq X$ a open set, $x \in U$ and $f: U \rightarrow Y$ a function then the following are equivalent*

1. *f is differentiable at x with differential $Df(x)$*
2. *There exists a open set $W \subseteq U$, $x \in W$, a linear function $Df(x): X \rightarrow Y$ and a ε-mapping $\varepsilon: W_x \rightarrow Y$ such that $\forall h \in W_x$ we have $f(x+h) - f(x) - Df(x)(h) = \varepsilon(h) \cdot \|h\|_X$*
3. *There exists a open set $W \subseteq U$, $x \in W$, a linear function $Df(x): X \rightarrow Y$ such that $\forall \varepsilon > 0$ there exists a $\delta > 0$ such that $\forall h \in W_x$ with $0 < \|h\| < \varepsilon$ we have $\frac{\|f(x+h) - f(x) - Df(x)(h)\|_Y}{\|h\|_X} < \varepsilon$*
4. *There exists a open set $W \subseteq U$, $x \in W$, a linear function $Df(x): X \rightarrow Y$ such that $\forall \{h_i\}_{i \in \mathbb{N}}$ with $\lim_{h \rightarrow 0} h_i = 0$ and $\forall i \in \mathbb{N}$ $\|h_i\| > 0$ and $h_i \in W_x$ we have that $\lim_{n \rightarrow \infty} \frac{\|f(x+h_i) - f(x) - Df(x)(h_i)\|_Y}{\|h_i\|_X} = 0$*

Let's study now some trival examples of functions differentiable at a point.

Example 14.14. (Differential of a constant function) Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be normed spaces, $y \in Y$, $U \subseteq X$ open and $C_{U,y}: U \rightarrow y$ defined by $C_{U,y}(x) = y$ then $C_{U,y}: U \rightarrow V$ is differentiable at every $x \in U$ and $DC_{U,y}(x) = C_{X,0}$ where $C_{X,0}$: $X \rightarrow Y$ is defined by $C_{X,0}(x) = 0$

Proof. Let $x \in U$ and take $\varepsilon > 0$ then for $h \in U_x$ with $\|h\|_X < 1$ we have $\|C_{U,y}(x + h) - C_{U,y}(x) - C_{X,0}(h)\|_Y = \|y - y - 0\|_Y = \|0\|_Y = 0 \leqslant \varepsilon \cdot \|h\|_X$. Also as $\forall x \in X$ we have $\|C_{X,0}(x)\|_Y = \|0\|_Y = 0 \leqslant 0 \cdot \|x\|_X$ it follows that $C_{X,0} \in L(X, Y)$. □

Example 14.15. Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces then $\forall U$ open in X we have that if $L\in L(X,Y)$ then $L: X\rightarrow Y$ is differentiable at every $x\in U$ with differential $DL(x)=L$

Proof. If $\varepsilon>0$ and $h\in X_x=X$ with $\|h\|_X<1$ then $\|L(x+h)-L(x)-L(h)\|_Y=\|L(x)+L(h)-L(x)-L(h)\|_Y=\|0\|_Y=0\leqslant\varepsilon\cdot\|h\|_X$ proving that L is differentiable at every $x\in U$ with $DL(x)=L$ (where we have used the fact that L is continuous as $L\in L(X,Y)$). □

14.1.2 Derivative of a function

Definition 14.16. *Given $\langle\mathbb{K},||\rangle$ (where $\mathbb{K}=\mathbb{R}$ or $\mathbb{C}$), $\langle Y,\|\|\rangle$ a normed space, $U\subseteq\mathbb{K}$ a open set, $x\in U$ and a function $f: U\rightarrow V$ then f has a* ***derivative*** *at x noted by $f'(x)$ if and only if $\forall\varepsilon>0$ there exists a $\delta>0$ such that $\forall h\in U_x$ with $0<|h|<\delta$ we have $\left\|\frac{f(x+h)-f(x)}{h}-f'(x)\right\|\leqslant\varepsilon$*

The above notation suggests that the derivative of a function, if it exists, is unique. The following theorem shows that this is indeed so and also the relation with the differential of a function.

Theorem 14.17. *Given $\langle\mathbb{K},||\rangle$ (where $\mathbb{K}=\mathbb{R}$ or $\mathbb{C}$), $\langle Y,\|\|\rangle$ a normed space, $U\subseteq\mathbb{K}$ a open set, $x\in U$ a open set and a function $f: U\rightarrow V$ then f is differentiable at x with differential $Df(x)$ if and only if f has a derivative $f'(x)$ at x. Furthermore if f is differentiable (or equivalent has a derivative) then $Df(x):\mathbb{K}\rightarrow Y$ is defined by $h\rightarrow Df(x)(h)=f'(x)\cdot h$ which means that $f'(x)=Df(x)(1)$. Note that as the differential is unique this theorem also proves that the derivative must be unique.*

Proof.

⇒. If f is differentiable with differential $Df(x)$, take then $f'(x)=Df(x)(1)$ then as $Df(x)\in L(\mathbb{K},Y)$ we have $\forall h\in\mathbb{K}$ that $Df(x)(h)=Df(x)(h\cdot 1)=h\cdot Df(x)(1)=h\cdot f'(x)$. Using 14.8 (3) we have that given a $\varepsilon>0$ there exists a $\delta>0$ such that $\forall h\in U_x$ with $0<|h|<\delta$ we have $\frac{\|f(x+h)-f(x)-Df(x)(h)\|}{|h|}\leqslant\varepsilon\Rightarrow\left\|\frac{f(x+h)-f(x)-f'(x)\cdot h}{h}\right\|\leqslant\varepsilon\Rightarrow\left\|\frac{f(x+h)-f(x)}{h}-f'(x)\right\|\leqslant\varepsilon$ proving that $f'(x)=Df(x)(1)$ is indeed a derivative of f at x and that $Df(x)(h)=f'(x)\cdot h$

⇐. Define $Df(x):\mathbb{K}\rightarrow Y$ by $h\rightarrow f'(x)\cdot h$ where $f'(x)$ is a derivative of f, which is trivially linear and is also continuous (as $\|Df(x)(h)\|_Y=\|f'(x)\cdot h\|_Y=\|f'(x)\|\cdot|h|$). Given $\varepsilon>0$ there exists a $\delta>0$ such that $\forall h\in U_x$ with $0<|h|<\delta$ we have $\left\|\frac{f(x+h)-f(x)}{h}-f'(x)\right\|\leqslant\varepsilon\Rightarrow\frac{\|f(x+h)-f(x)-Df(x)(h)\|}{|h|}=\left\|\frac{f(x+h)-f(x)-f'(x)\cdot h}{h}\right\|=\left\|\frac{f(x+h)-f(x)}{h}-f'(x)\right\|\leqslant\varepsilon$ which proves by 14.8 (3) that f is differentiable at x with $Df(x)(h)=f'(x)\cdot h$. As the differential is unique and $Df(x)(1)=f'(x)$ the derivative must be unique. □

As the differentiability of a function is a local property the above theorem shows that this is also true for the existance of the derivative of a function.

Corollary 14.18. *Given $\langle \mathbb{K}, || \rangle$ (where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$), $\langle Y, \|\| \rangle$ a normed space, $U \subseteq \mathbb{K}$, $x \in U$ a open set, a function $f: U \rightarrow V$ and $W \subseteq U$ a open set with $x \in W$ then f has a derivative at x if and only of $f_{|W}$ has a derivative at x and $(f_{|W})'(x) = f'(x)$*

Proof. Thius follows from 14.12 and the above theorem. □

14.1.3 Properties of the differential

Definition 14.19. *Given a finite family $\{X_i\}_{i \in \{1, \ldots, n\}}$ of vector spaces over the field $\mathbb{K}$ then given a $i \in \{1, \ldots, n\}$ and $x \in \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} X_j$ we define $(x|i): X_i \rightarrow \prod_{i \in \{1, \ldots, n\}} X_i$ by $t \rightarrow (i \rightarrow x)(t)$ where $\forall j \in \{1, \ldots, n\}$ $((i \rightarrow x)(t))_j = \left\{ \begin{array}{l} x_j \text{ if } j \in \{1, \ldots, n\} \setminus \{i\} \\ t \text{ if } j = i \end{array} \right.$.*

Remark 14.20. Essentially $(i \rightarrow x)(t)$ substitue the i-the element of x by t, so we have that $(i \rightarrow (x_1, \ldots, x_{i1}, x_i, x_{i+1}, \ldots, x_n))(t) = (x_1, \ldots, x_{i-1}, t, x_i, \ldots, x_n)$

Proposition 14.21. *Let $\{X_i\}_{i \in \{1, \ldots, n\}}$ be a finite family of vector spaces over the field $\mathbb{K}$, $x \in \prod_{i \in \{1, \ldots, n\}} X_i$ then $\forall i \in \{1, \ldots, n\}$ we have*

1. $(i \rightarrow x)(x_i) = x$
2. $(i \rightarrow x)(t) - (i \rightarrow x)(s) = (i \rightarrow 0)(t - s)$

Proof.

1. Let $k \in \{1, \ldots, n\}$ then $((i \rightarrow x)(x_i))_k = \left\{ \begin{array}{l} x_k \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \\ x_i \text{ if } k = i \end{array} \right. = x_k$ so that $(i \rightarrow x)(x_i) = x$
2. Let $k \in \{1, \ldots, n\}$ then

$$\begin{array}{rcl} ((i \rightarrow x)(t) - (i \rightarrow x)(s))_k & = & ((i \rightarrow x)(t))_k - ((i \rightarrow x)(s)) \\ & = & \left\{ \begin{array}{l} x_k \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \\ t \text{ if } k = i \end{array} \right. - \\ & & \left\{ \begin{array}{l} x_k \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \\ s \text{ if } k = 1 \end{array} \right. \\ & = & \left\{ \begin{array}{l} 0_i \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \\ t - s \text{ if } k = i \end{array} \right. \\ & = & (i \rightarrow 0)(s - t) \end{array}$$

□

Theorem 14.22. *Given a finite family $\{X_i\}_{i \in \{1, \ldots, n\}}$ of vector spaces over the field $\mathbb{K}$ then given a $i \in \{1, \ldots, n\}$ and $0 \in \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} X_j$, where $0_j \in X_j$ is the neutral element of X_j then*

1. *$\forall i, k \in \{1, \ldots, n\}$ and $x \in X_i$* then $((i \rightarrow 0)(x))_k = \delta_{i,k} \cdot x$

2. $\forall i \in \{1, \ldots, n\}$ *we have that* $(i \rightarrow 0) \in L\left(X_i, \prod_{i \in \{1, \ldots, n\}} X_i\right)$ *[using the maximum norm* $\|\|$ *on* $\prod_{i \in \{1, \ldots, n\}} X_i$*]*

Proof.

1. Let $i \in \{1, \ldots, n\}$ then for $k \in \{1, \ldots, n\}$ we have either

   $\boldsymbol{k \neq i}$**.** then $((i \rightarrow 0)(x)) = \begin{cases} 0_k \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \\ x \text{ if } k = i \end{cases} = 0_k = \delta_{i,k} \cdot x$

   $\boldsymbol{k = i}$**.** then $((i \rightarrow 0)(x)) = \begin{cases} 0_k \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \\ x \text{ if } k = i \end{cases} = x = \delta_{i,k} \cdot x$

2. We must prove that $(i \rightarrow 0)$ is linear and continuous. Let $i \in \{1, \ldots, n\}$ then

   a. Let $\alpha, \beta \in \mathbb{K}$ and $x, y \in X_i$ then given $k \in \{1, \ldots, n\}$ we have

   $$\begin{aligned} ((i \rightarrow 0)(\alpha \cdot x + \beta \cdot y))_k &\underset{(1)}{=} \delta_{i,k} \cdot (\alpha \cdot x + \beta \cdot y) \\ &= \alpha \cdot \delta_{i,k} \cdot x + \beta \cdot \delta_{i,k} \cdot y \\ &= \alpha \cdot ((i \rightarrow 0)(x))_k + \beta \cdot ((i \rightarrow 0)(y))_k \\ &= (\alpha \cdot (i \rightarrow 0)(x) + \beta \cdot (i \rightarrow 0)(y))_k \end{aligned}$$

   proving that $(i \rightarrow 0)(\alpha \cdot x + \beta \cdot y) = \alpha \cdot (i \rightarrow 0)(x) + \beta \cdot (i \rightarrow 0)(y)$. So $(i \rightarrow 0)$ is linear.

   b. Let $x \in X_i$ then we have

   $$\begin{aligned} \|(i \rightarrow 0)(x)\| &= \max\left(\{\|((i \rightarrow 0)(x))_j\|_{j \in \{1, \ldots, n\}} | j \in \{1, \ldots, n\}\}\right) \\ &\underset{(1)}{=} \max\left(\{\|\delta_{i,j} \cdot x\|_j | j \in \{1, \ldots, n\}\}\right) \\ &= \|x\|_j \end{aligned}$$

   which by 12.175 proves that $(i \rightarrow 0)$ is continuous.

□

Theorem 14.23. *Given a finite family* $\{X_i\}_{i \in \{1, \ldots, n\}}$ *of vector spaces over the field* $\mathbb{K}$ *then for every* $x \in \prod_{i \in \{1, \ldots, n\}} X_i$ *we have* $x = \sum_{i=1}^{n} (i \rightarrow 0)(x_i)$

Proof. Let $k \in \{1, \ldots, n\}$ then

$$\begin{aligned} \left(\sum_{i=1}^{n} (i \rightarrow 0)(x_i)\right)_k &\underset{10.10}{=} \sum_{i=1}^{n} ((i \rightarrow 0)(x_i))_k \\ &\underset{14.22}{=} \sum_{i=1}^{n} (\delta_{i,k} \cdot x_i) \\ &= x_k \end{aligned}$$

proving that $\sum_{i=1}^{n} (i \rightarrow 0)(x_i) = x$. □

Theorem 14.24. *Given a finite family* $\{\langle X_i, \|\|_i \rangle\}_{i \in \{1, \ldots, n\}}$ *of vector spaces over the field* $\mathbb{K}$*,* $X = \prod_{i \in \{1, \ldots, n\}} X_i$ *the product space together with the maximum norm* $\|\|$ *(see 12.79) (so that* $\langle X, \|\| \rangle$ *forms a normed vector space over* $\mathbb{K}$*) then given* $i \in \{1, \ldots, n\}$*,* $t \in X_i$ *and* $x \in \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} X_j$ *we have that* $(i \rightarrow x): X_i \rightarrow \prod_{j \in \{1, \ldots, n\}} X_j$ *is differentiable at* t *with* $D((i \rightarrow x)) = (i \rightarrow 0)$*.*

Proof. Let $i\in\{1,\ldots,n\}$ then using 14.22 we have that

$$(i\rightarrow 0)\in L\left(X_i, \prod_{j\in\{1,\ldots,n\}} X_j\right)$$

Take now $\varepsilon>0$, $t\in X_i$ then if $h\in (X_i)_t=X_i$ with $\|h\|_i<1$ we have using 14.21 that $(i\rightarrow x)(t+h)-(i\rightarrow x)(t)=(i\rightarrow 0)(h)$. Hence $\|(i\rightarrow x)(t+h)-(i\rightarrow x)(t)-(i\rightarrow 0)(h)\|=\|0\|=0\leqslant\varepsilon$ proving that $(i\rightarrow x)$ is differentiable at t with differential $(i\rightarrow 0)$. □

Note 14.25. Given a finite family $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ of vector spaces over the field $\mathbb{K}$, $X=\prod_{i\in\{1,\ldots,n\}} X_i$ the product space together with the maximum norm $\|\|$ (see 12.79), then $\forall x\in\prod_{i\in\{1,\ldots,n\}} X_i$ we have by the above and 14.10 (functions that are differentiable at a point are coninuous at that point) that $\forall i\in\{1,\ldots,n\}$ $(i\rightarrow x)$ is continous at X_i. So given a open set $U\subseteq\prod_{i\in\{1,\ldots,n\}} X_i$ we have that $\{t\in X_i|(i\rightarrow x)(t)\in U\}=(i\rightarrow x)^{-1}(U)$ is a open set in X^i.

Theorem 14.26. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces over $\mathbb{K}$, $U\subseteq X$ a open set and $f: U\rightarrow Y$, $g: U\rightarrow Y$ be functions differentiable at $x\in U$ then*

1. *$f+g: U\rightarrow Y$ is differentiable at x with $D(f+g)(x)=Df(x)+Dg(x)$*
2. *$\forall\alpha\in\mathbb{K}$ then $\alpha\cdot f: U\rightarrow Y$ is differentiable at x with $D(\alpha\cdot f)(x)=\alpha\cdot Df(x)$*

Proof.

1. Let $\varepsilon>0$ then as f,g are differentiable at x there exists $\delta_1,\delta_2>0$ such that if $h\in U_x$ we have if $\|h\|_X<\delta_1$ then $\|f(x+h)-f(x)-Df(x)(h)\|_Y\leqslant\frac{\varepsilon}{2}\cdot\|h\|_X$ and if $\|h\|_X<\delta_2$ then $\|g(x+h)-g(x)-Dg(x)(h)\|_Y\leqslant\frac{\varepsilon}{2}\cdot\|h\|_X$. So if $h\in U_x$ and $\|h\|_X<\delta=\min(\delta_1,\delta_2)$ we have that $\|(f+g)(x+h)-(f+g)(h)-(Df(x)+Dg(x))(h)\|_Y=\|f(x+h)-f(x)-Df(x)(h)+g(x+h)-g(x)-Dg(x)(h)\|_Y\leqslant\|f(x+h)-f(x)-Df(x)(h)\|_Y+\|g(x+h)-g(x)-Dg(x)(h)\|_Y\leqslant\frac{\varepsilon}{2}\cdot\|h\|_X+\frac{\varepsilon}{2}\cdot\|h\|_X=\varepsilon\cdot\|h\|_X$ proving that $f+g$ s differentiable with differential $Df(x)+Dg(x)$.

2. Let $\varepsilon>0$ then by the differentiability of f at x there exists a $\delta>0$ such that if $h\in U_x$ and $\|h\|_X<\delta$ we have $\|f(x+h)-f(x)=Df(x)(h)\|_Y\leqslant\frac{\varepsilon}{|\alpha|+1}\cdot\|h\|_X$. So $\|(\alpha\cdot f)(x+h)-(\alpha\cdot f)(h)-\alpha\cdot Df(x)(h)\|_Y=|\alpha|\cdot\|f(x+h)-f(x)-Df(x)(h)\|_Y\leqslant|\alpha|\cdot\frac{\varepsilon}{|\alpha|+1}\cdot\|h\|_X<\varepsilon\cdot\|h\|_X$. Proving that $\alpha\cdot f$ is indeed differentiable at $x\in U$ with differential $\alpha\cdot Df(x)$. □

Theorem 14.27. (The chain rule) *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$,$\langle Z,\|\|_X\rangle$ be normed spaces, $U\subseteq X$*

, $V\subseteq Y$ be open sets, $f: U\rightarrow X$, $g: V\rightarrow Y$ functions, $x\in U\bigcap f^{-1}(V)$ such that f is differentiable at x and g is differentiable at $f(x)$ then $g\circ f: U\bigcap f^{-1}(V)\rightarrow Z$ is differentiable at x with $D(g\circ f)(x)=Dg(f(x))\circ Df(x)$

Proof. First as $f: U \to Y$ is differentiable at x we have by 14.10 that f is continuous at x so that $f^{-1}(V)$ is open proving that $U \bigcap f^{-1}(V)$ is open (needed for differentiability to make sense) Using the alternative definition for differentiability (see 14.13) there exists ε-mapping $\varepsilon_f: U_x \to Y$ and $\varepsilon_g: V_{f(x)} \to Y$ such that

$$\forall h \in \left(U \bigcap f^{-1}(V)\right)_x \text{ we have } f(x+h) - f(x) - Df(x)(h) = \varepsilon_f(h) \cdot \|h\|_X \tag{14.12}$$

and

$$\forall k \in V_{f(x)} \text{ we have } g(f(x)+k) - g(f(x)) - Dg(f(x))(k) = \varepsilon_g(k) \cdot \|k\|_Y \tag{14.13}$$

If $h \in (U \bigcap f^{-1}(V))_x$ then $x+h \in U \bigcap f^{-1}(V)$ so that $(f(x+h) - f(x)) + f(x) = f(x+h) \in V$ proving that

$$\forall h \in \left(U \bigcap f^{-1}(V)\right)_x \text{ we have that } f(x+h) - f(x) \in V_{f(x)} \tag{14.14}$$

Combining the above with 14.13 gives then that for $h \in (U \bigcap f^{-1}(V))_x$ we have

$$\begin{aligned}
\varepsilon_g(f(x+h) - f(x)) \cdot \|f(x+h) - f(x)\|_Y &= g(f(x) + f(x+h) - f(x)) - g(f(x)) - Dg(f(x))(f(x+h) - f(x)) \\
&= g(f(x+h)) - g(f(x)) - Dg(f(x))(f(x+h) - f(x)) \\
&\underset{14.12}{=} g(f(x+h)) - g(f(x)) - Dg(f(x))(\varepsilon_f(h) \cdot \|h\|_X + Df(x)(h)) \\
&= g(f(x+h)) - g(f(x)) - D\, g(f(x))(\varepsilon_f(h) \cdot \|h\|_X) - Dg(f(x))(Df(x)(h)) \\
&= (g \circ f)(x+h) - (g \circ f)(x) - D\, g(f(x))(\varepsilon_f(h) \cdot \|h\|_X) - Dg(f(x))(Df(x)(h))
\end{aligned}$$

so that $(g \circ f)(x+h) - (g \circ f)(x) - Dg(f(x))(Df(x)(h)) = \varepsilon_g(f(x+h) - f(x)) \cdot \|f(x+h) - f(x)\|_Y + Dg(f(x))(\varepsilon_f(h) \cdot \|h\|_X)$ hence

$$(g \circ f)(x+h) - (g \circ f)(x) - Dg(f(x)) \circ Df(x)(h) = \varepsilon_g(f(x+h) - f(x)) \cdot \|f(x+h) - f(x)\|_Y + Dg(f(x))(\varepsilon_f(h) \cdot \|h\|_X) \tag{14.15}$$

Define now $\zeta: (U \bigcap f^{-1}(V))_x \to Z$ by

$$\zeta(h) = \begin{cases} \varepsilon_g(f(x+h) - f(x)) \cdot \frac{\|f(x+h) - f(x)\|_Y}{\|h\|_X} + Dg(f(x))(\varepsilon_f(h)) \text{ if } h \neq 0 \\ 0 \text{ if } h = 0 \end{cases}$$

then as for $h = 0$ we have that $(g \circ f)(x+h) - (g \circ f)(x) - (Dg(f(x)) \circ Df(x))(h) = (g \circ f)(x) - (g \circ f)(x) - (Dg(f(x)) \circ Df(x))(0) = 0$ we have by 14.15 that

$$(g \circ f)(x+h) - (g \circ f)(x) - Dg(f(x)) \circ Df(x)(h) = \zeta(h) \cdot \|h\|_X \tag{14.16}$$

If we prove that ζ is continuous at 0 and thus a ε-mapping we have by 14.13 and the fact that $Dg(f(x))\circ Df(x)\in L(X,Z)$ (see 12.193) that $g\circ f: U\bigcap f^{-1}(V)\rightarrow Z$ is differentiable at x with differential $Dg(f(x))\circ Df(x)$.

To prove continuity of ζ take $\varepsilon>0$. By the continuity of ε_f at 0 there exists a $\delta_1>0$ such that if $\|h\|_X<\delta_1$ then $\|\varepsilon_f(h)\|<1$ and using 14.12 we have if h is also a element of $(U\bigcap f^{-1}(V))_x$ that $\|f(x+h)-f(x)\|_Y\leqslant\|\varepsilon_f(h)\cdot\|h\|_X+Df(x)(h)\|_Y\leqslant\|\varepsilon_f(h)\|\cdot\|h\|_X+\|Df(x)\|\cdot\|h\|_X<\|h\|_X+\|Df(x)\|\cdot\|h\|_X$ so that

$$\forall h\in\left(U\bigcap f^{-1}(V)\right)_x \text{ with } 0<\|h\|_X<\delta_1 \text{ we have } \frac{\|f(x+h)-f(x)\|_Y}{\|h\|_X}<(1+\|Df(x)\|) \tag{14.17}$$

By continuity of ε_g at 0 there exists a $\delta_2>0$

$$\forall k\in V_{f(x)} \text{ with } \|k\|_Y<\delta_2 \text{ we have } \|\varepsilon_g(k)\|_Z<\frac{\varepsilon}{2\cdot(1+\|Df(x)\|)} \tag{14.18}$$

As f is differentiable at x and thus by 14.10 continuous at x there exists a $\delta_3>0$ such that if $h\in(U\bigcap f^{-1}(V))_x \underset{14.14}{\Rightarrow} f(x+h)-f(x)\in V_{f(x)}$ and $\|h\|_X<\delta_3$ then $\|f(x+h)-f(x)\|_Y<\delta_2 \underset{14.17}{\Rightarrow} \|\varepsilon_g(f(x+h)-f(x))\|_Z<\frac{\varepsilon}{2\cdot(1+\|Df(x)\|)}$. Using this with 14.17 gives

$$\text{If } h\in\left(U\bigcap f^{-1}(V)\right)_x \text{ with } 0<\|h\|_X<\min(\delta_1,\delta_3) \text{ we have } \|\varepsilon_g(f(x+h)-f(x))\|_X\cdot\frac{\|f(x+h)-f(x)\|_Y}{\|h\|_X}<\frac{\varepsilon}{2} \tag{14.19}$$

As $Dg(f(x))\in L(Y,Z)$ we have that $\|Dg(f(x))(\varepsilon_f(h))\|_Z\leqslant\|Dg(f(x))\|\cdot\|\varepsilon_f(h)\|_Y$ and as ε_f is continuous at 0 there exists a δ_4 such that if $\|h\|_X<\delta_4$ then $\|\varepsilon_f(h)\|_Y<\frac{\varepsilon}{2\cdot(1+\|Dg(f(x))\|)}$ and thus

$$\text{If } \|h\|_X<\delta_4 \text{ then } \|Dg(f(x))(\varepsilon_f(h))\|_Z<\frac{\varepsilon}{2} \tag{14.20}$$

Finally if we take $\delta=\min(\delta_1,\delta_3,\delta_4)$ we have for $h\in(U\bigcap f^{-1}(V))_x$ with $\|h\|_X<\delta$ that either

$h=0$. then $\|\zeta(h)-0\|_Z=\|0-0\|_X=0<\varepsilon$

$h\neq 0$. then $0<\|h\|_X<\delta$ and thus $\|\zeta(h)-0\|_Z=\|\zeta(h)\|_Z=\left\|\varepsilon_g(f(x+h)-f(x))\cdot\frac{\|f(x+h)-f(x)\|_Y}{\|h\|_X}+Dg(f(x))(\varepsilon_f(h))\right\|_Z\leqslant\left\|\varepsilon_g(f(x+h)-f(x))\cdot\frac{\|f(x+h)-f(x)\|_Y}{\|h\|_X}\right\|_Z+\|Dg(f(x))(\varepsilon_f(h))\|_Z<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$ by 14.19 and 14.20.

So in all cases we have that $\|\zeta(h)-0\|_Z<\varepsilon$ proving continuity of ζ at 0 and thus that ζ is a ε-mapping. $\square$

The above theorem is a generic form of the Chain Rule more know in the form below.

Corollary 14.28. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$, $\langle Z, \|_Z\|\rangle$ be normed spaces over $\mathbb{K}$, $U \subseteq X$ a open set, $V \subseteq Y$ a open set, $f: U \rightarrow Y$, $g: V \rightarrow Z$ functions with $f(U) \subseteq V$, such that f is differentiable at $x \in U$ and g is differentiable at $f(x) \in W$ then $g \circ f: U \rightarrow W$ is differentiable at x with $D(g \circ f)(x) = Dg(f(x)) \circ Df(x)$*

Proof. As $f(U) \subseteq V$ we have that $U \subseteq f^{-1}(V)$ proving that $U = U \bigcap f^{-1}(V)$, using then the previous theorem we have that $g \circ f: U \rightarrow Z$ is differentiable at $x \in U$ with $D(g \circ f)(x) = Df(g(x)) \circ Df(x)$ □

The chain rule can be translated to the concept of derivates

Corollary 14.29. *Let $\mathbb{K}$ be the field $\mathbb{R}$ or $\mathbb{C}$, U, V open sets in $\mathbb{K}$, $f: U \rightarrow \mathbb{K}$, $g: V \rightarrow \mathbb{R}$ functions such that for $x \in U \bigcap f^{-1}(V)$ f has a derivate on x and g has a derivative on $f(x)$ then $g \circ f$ has a derivate on x and $(g \circ f)'(x) = g'(f(x)) \cdot f'(x)$*

Proof. Using 14.17 we have that f is differentiable at x with $f'(x) = Df(x)(1)$ and g is differentiable at $f(x)$ with $Dg(x)(1)$. By the chain rule (see 14.27) we have that $g \circ f$ is differentiable at x with $D(g \circ f)(x) = Dg(f(x)) \circ Df(x)$. Using then 14.17 again we have that $g \circ f$ has a derivate $(g \circ f)'(x)$ at x with

$$\begin{aligned}
(g \circ f)'(x) &= (Dg(f(x)) \circ Df(x))(1) \\
&= (Dg(f(x)))(Df(x)(1)) \\
&= (Dg(f(x)))(f'(x)) \\
&\underset{14.17}{=} g'(f(x)) \cdot f'(x)
\end{aligned}$$

□

14.1.4 Partial differentiation

Definition 14.30. *Let $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1, \ldots, n\}}$ be a finite family of normed spaces, $\langle Y, \|\|_Y\rangle$ a normed space, $X = \left\langle \prod_{i \in \{1, \ldots, n\}} X_i, \|\|\right\rangle$ the normed space equipped with the maximum norm (see 12.79), $U \subseteq \prod_{i \in \{1, \ldots, n\}} X_i$ a open set, $x \in U$ and $f: U \rightarrow Y$ a function, $i \in \{1, \ldots, n\}$, then $f: U \rightarrow Y$ is i-partial differentiable at x with i-partial differential $D_i f(x) \in L(X_i, Y)$ if $f \circ (i \rightarrow x): (i \rightarrow x)^{-1}(U) \rightarrow V$ is differentiable at x_i with $D_i f(x) = D(f \circ (i \rightarrow x))(x_i)$ [This make sense as by 14.25 $(i \rightarrow x)^{-1}(U)$ is open in X_i and we have trivially $x_i \in (i \rightarrow x)^{-1}(U)$ as $(i \rightarrow x)(x_i) \underset{14.21}{=} x \in U$].*

The next theorem shows that the partial differentials exists if a function is differentiable.

Theorem 14.31. *Let $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1, \ldots, n\}}$ be a finite family of normed spaces, $\langle Y, \|\|_Y\rangle$ a normed space, $\left\langle \prod_{i \in \{1, \ldots, n\}} X_i, \|\|\right\rangle$ the normed space equipped with the maximum norm (see 12.79), $U \subseteq \prod_{i \in \{1, \ldots, n\}} X_i$ a open set, $x \in U$ and $f: U \rightarrow Y$ differentiable at $x \in U$ then*

1. *$\forall i \in \{1, \ldots, n\}$ f is i-partial differentiable at $x_i \in X_i$ with $D_i f(x) = Df(x) \circ (i \rightarrow 0)$ [or using a less formal notation $(D_i f(x))(t) = Df(x)(0, \ldots, t, \ldots, 0)$]*

2. $Df(x) = \sum_{i=1}^{n} D_i f(x) \circ \pi_i$ *[or* $\forall n \in \mathbb{N}$ *we have* $Df(x)(h) = \sum_{i=1}^{n} D_i f(x)(h_i)$

Proof.

1. Using 14.24 and 14.12 we have that $(i \to x): X_i \to \prod_{i \in \{1,\ldots,n\}} X_i$ is differentiable at x_i with $D(i \to x)(x_i) = (i \to 0)$. Using the chain rule 14.27 we have that $f \circ (i \to x): (i \to x)^{-1}(U) \to Y$ is differentiable at x_i with $D(f \circ (i \to x)) = Df((i \to x)(x_i)) \circ (i \to 0) \underset{14.21}{=} Df(x) \circ (i \to 0)$

2. Let $h \in X$ then we have

$$\begin{aligned}\sum_{i=1}^{n} D_i f(x)(h_i) &= \sum_{i=1}^{n} Df(x)((i \to 0)(h_i)) \\ &\underset{Df(x)\,\text{is linear}}{=} Df(x)\left(\sum_{i=1}^{n} (i \to 0)(h_i)\right) \\ &\underset{14.23}{=} Df(x)(h)\end{aligned}$$

$\square$

Just like we have defined the derivative of a function, we define now the partial derivative of a function.

Definition 14.32. *Let* $\langle \mathbb{K}^n, \|\|\rangle$ *be the vector space* $\mathbb{K}^n$ *equipped with the maximum norm,* $\langle Y, \|\|_Y\rangle$ *a normed space,* $U \subseteq \mathbb{K}^n$ *a open set,* $x \in U$*,* $i \in \{1, \ldots, n\}$ *then* f *has a* ***i-partial derivative at*** x *noted by* $\partial_i f(x) \in Y$ *if and only if* $\forall \varepsilon > 0$ *there exists a* $\delta > 0$ *such that if* $0 < |h| < \delta$ *and* $h \in ((i \to x)^{-1}(U))_x$ *then* $\left\|\frac{f((i \to x)(x_i + h)) - f(x)}{h} - \partial_i f(x)\right\|_Y < \varepsilon$ *[or a less formal notation* $\left\|\frac{f(x_1, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_n)}{h}\right\| < \varepsilon$*]*

The above definition suggest that the partial derivate is unique wich is indeed true as is expressed in the following theorem which also show the relation with the partial differential.

Theorem 14.33. *Let* $\langle \mathbb{K}^n, \|\|\rangle$ *be the vector space* $\mathbb{K}^n$ *equipped with the maximum norm,* $\langle Y, \|\|_Y\rangle$ *a normed space,* $U \subseteq \mathbb{K}^n$ *a open set,* $x \in U$*,* $i \in \{1, \ldots, n\}$ *then* f *has a i-partial derivative at* x *if and only if* f *has a i-partial differential. Further if the i-partial derivative exists (or the i-partial differential exists) at* x *then* $\partial_i f(x) = D_i f(x)(1)$ *so that* $D_i f(x)(h) = D_i f(x)(h \cdot 1) = h \cdot D_i f(x)(1)$ *Note as the i-partial differential is unique this proves also that the i-partial derivative is unique.*

Proof.

⇒. Assume that $\partial_i f(x)$ exist.Then given $\varepsilon > 0$ there exists a $\delta > 0$ such that $h \in ((i \to x)^{-1}(U))_x$ and $0 < |h| < \delta$ then we have that

$$\left\|\frac{f((i \to x)(x_i + h) - f(x)}{h} - \partial_i f(x)\right\|_Y < \varepsilon$$

As $\left\|\frac{f((i\to x)(x_i+h))-f((i\to x)(x_i))}{h}-\partial_i f(x)\right\|_Y \underset{14.21}{=} \left\|\frac{f((i\to x)(x_i+h))-f(x)}{h}-\partial_i f(x)\right\|_Y < \varepsilon$ proving that the derivative of $f\circ(i\to x)$ at x_i exists and $(f\circ(i\to x))'(x_i)=\partial_i f(x)$. Using 14.17 we have that $D(f\circ(i\to x))(x_i)$ exists and $D(f\circ(i\to x))(x_i)(1)=(f\circ(i\to x))'(x_i)=\partial_i f(x)$. Using the definition of the partial differential (see 14.30) it follows that

$$D_i f(x) \text{ exists and } D_i f(x)(1)$$

$\Leftarrow$. Assume that $D_i f(x)$ exists at x then by definition $D(f\circ(i-x))(x_i)$ exists. Using 14.17 we have that $(f\circ(i\to x))'(x_i)$ exists so that given $\varepsilon>0$ there exists a $\delta>0$ such that $\forall h\in((i\to x)^{-1}(U))_x$ with $0<|h|<\delta$ we have

$$\left\|\frac{(f\circ(i\to x))(x_i+h)-(f\circ(i\to x))(x_i)}{h}-(f\circ(i\to x))'(x_i)\right\|_Y<\varepsilon$$

which as $\left\|\frac{(f\circ(i\to x))(x_i+h)-(f\circ(i\to x))(x_i)}{h}-(f\circ(i\to x))'(x_i)\right\|_Y = \left\|\frac{f((i\to x)(x_i+h))-f((i\to x)(x))}{h}-(f\circ(i\to x))'(x_i)\right\|_Y \underset{14.21}{=} \left\|\frac{(f\circ(i\to x))(x_i+h)-f(x)}{h}-(f\circ(i\to x))'(x_i)\right\|_Y$ proves that $\partial_i f(x)$ exists. □

We have then the equivalence of theorem 14.31

Theorem 14.34. *Let $\langle\mathbb{K}^n,\|\|\rangle$ be the vector space $\mathbb{K}^n$ equipped with the maximum norm based on the absolute value norm, $\langle Y,\|\|_Y\rangle$ a normed space, $U\subseteq\mathbb{K}^n$ a open set, $x\in U$ and $f:U\to Y$ a function that is differentiable at x then*

1. *$\forall i\in\{1,\ldots,n\}$ f has a partial derivative $\partial_i f(x)$ and the partial differential $D_i f(x)\in L(\mathbb{K},Y)$ is defined by $k\to k\cdot\partial_i f(x)$ (so $D_i f(x)(1)=\partial_i f(x)$)*
2. *If $\{e_i\}_{i\in\{1,\ldots,n\}}$ is the canonical base on $\mathbb{K}^n$ (defined by $(e_i)_j=\delta_{i,j}$) then $\partial_i f(x)=Df(x)(e_i)$ and $Df(x)(h)=\sum_{i\in\{1,\ldots,n\}} h_i\cdot Df(x)(e_i)=\sum_{i\in\{1,\ldots,n\}} h_i\cdot\partial_i f(x)$. So if we use matrix notation and define $h=\begin{pmatrix} h_1\\ \cdot\\ \cdot\\ \cdot\\ h_n\end{pmatrix}$ and $\partial f(x)=(\partial_1 f(x),\ldots,\partial_n f(x))$ then we have $Df(x)(h)=\partial f(x)\cdot h$ (using the matrix product).*

Proof.

1. Let $i\in\{1,\ldots,n\}$ then by 14.31 $D_i f(x)$ exists so using 14.33 we have that $\partial_i f(x)$ exists and $\forall h\in\mathbb{K}$ $D_i f(x)(h)=h\cdot\partial_i f(x)$ or $\partial_i f(x)=D_i f(x)(1)$
2. Using 14.31 we have $\forall i\in\{1,\ldots,n\}$ that $D_i f(x)=Df(x)\circ(i\to 0)$. As $((i\to 0)(1))_k \underset{14.22}{=} \delta_{i,k}$ proving that $((i\to 0)(1))=e_i$ it follows thate $\partial_i f(x) \underset{(1)}{=} D_i f(x)(1)=Df(x)((i\to 0)(1))=Df(x)(e_i)$. So

$$\partial_i f(x)=Df(x)(e_i) \tag{14.21}$$

Finally as $h=\sum_{i=1}^{n} h_i\cdot e_i$ we have by the linearity of $Df(x)$ that $Df(x)(h)=Df(x)(\sum_{i=1}^{n} h_i\cdot e_i)=\sum_{i=1}^{n} h_i\cdot Df(x)(e_i) \underset{14.21}{=} \sum_{i=1}^{n} h_i\cdot\partial_i f(x)$ proving that

$$Df(x)(h)=\sum_{i=1}^{n} h_i\cdot\partial_i f(x)$$

□

Theorem 14.35. *Let $\langle X,\|\|_X\rangle$ be a normed space, $\{\langle Y_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a family of normed spaces and $\left\langle \prod_{i\in\{1,\ldots,n\}} Y_i,\|\|\right\rangle$ the normed space equipped with the maximum norm, $U\subseteq X$ open, $x\in U$ and $f\colon U\to Y$ a function $x\to f(x)=(f_1(x),\ldots,f_n(x))$ where $f_i=\pi_i\circ f\colon U\to Y_i$. Then f is differentiable at x if and only if $\forall i\in\{1,\ldots,n\}$ f_i is differentiable at x. If $Df(x)$ exists then $Df(x)(h)=(Df_1(x)(h),\ldots,Df_n(x)(h))$ or equivalently $\pi_i\circ Df(x)=Df_i(x)$.*

Proof.

- **⇒.** As f is differentiable, given $\varepsilon>0$ there exists a $\delta>0$ such that if $h\in U_x$ and $\|h\|_X<\delta$ then $\|f(x+h)-f(x)-Df(x)(h)\|\leqslant\|h\|_X\cdot\varepsilon$, by the definition of the maximum norm we have then $\max(\{\|f_i(x+h)-f_i(x)-\pi_i(Df(x)(h))\|_i\}|i\in\{1,\ldots,n\})\leqslant\varepsilon\cdot\|h\|_X$ so that $\forall i\in\{1,\ldots,n\}$ we have $\|f_i(x+h)-f_i(x)-(\pi_i\circ Df(x))(h)\|_i\leqslant\varepsilon\cdot\|h\|_X$ proving as π_i, $Df(x)$ is linear and continuous that $f_i\colon X\to Y_i$ is differentiable at x with differential $Df_i(x)=\pi_i\circ Df(x)$.
- **⇐.** Suppose that $\forall i\in\{1,\ldots,n\}$ $f_i\colon X\to Y_i$ is differentiable at x then given $\varepsilon>0$ there exists δ_i such that $\forall h\in U_x$ with $\|h\|_X<\delta_i$ $\|f_i(x+h)-f_i(x)-Df_i(x)(h)\|_i\leqslant\varepsilon\cdot\|h\|_X$ so if we define $Df(x)\colon X\to\prod_{i\in\{1,\ldots,n\}} Y_i$ by $h\to Df(x)=(Df_1(x)(h),\ldots,Df_n(x)(h))$ which is obviously linear as $Df_i(x)$ is linear. As $\|Df(x)(h)\|=\max(\{\|Df_i(x)(h)\|_i\}|i\in\{1,\ldots,n\})\leqslant\max(\{\|Df_i(x)\|\cdot\|h\|_X\}|i\in\{1,\ldots,n\})\leqslant\max(\{\|Df_i(x)\||i\in\{1,\ldots,n\}\})\cdot\|h\|_X$ $Df(x)$ is also continuous. Take now $\delta=\min(\{\delta_i|i\in\{1,\ldots,n\}\})$ then $\forall i\in\{1,\ldots,n\}$ we have if $h\in U_x$ and $\|h\|_X\leqslant\delta\Rightarrow\forall i\in\{1,\ldots,n\}$ $\|h\|_X\leqslant\delta_i\Rightarrow\|f_i(x+h)-f_i(x)-Df_i(x)(h)\|_i\leqslant\varepsilon\cdot\|h\|_X\Rightarrow\|f(x+h)-f(x)-Df(x)(h)\|=\max(\{\|f_i(x+h)-f_i(x)-Df_i(x)(h)\|_i|i\in\{1,\ldots,n\}\})\leqslant\varepsilon\cdot\|h\|_X$ proving that f is differentiable at x with differential $Df(x)=(Df_1(x),\ldots,Df_n(x))$. □

Theorem 14.36. (Jacobian matrix) *Let $\langle\mathbb{K}^n,\|\|_n\rangle$, $\langle\mathbb{K}^m,\|\|_m\rangle$ be two normed spaces based on $\langle\mathbb{K},||\rangle$ (with the maximum norm), a open set $U\subseteq\mathbb{K}^n$, $x\in U$ and $f\colon U\to V$ a differentiable function then $\forall j\in\{1,\ldots,m\}$ we have*

$$\pi_j(Df(x)(h))=(Df(x)(h))_j=Df_j(x)(h)=\sum_{i\in\{1,\ldots,n\}}\partial_i(\pi_j\circ f)(x)\cdot h_i=\sum_{i\in\{1,\ldots,n\}}\partial_i f_j(x)\cdot h_i.$$

If we define the matrix $\frac{\partial(f_1,\ldots,f_m)(x)}{\partial(1,\ldots,n)}=\begin{pmatrix}\partial_1 f_1(x) & \cdots & \partial_n f_1(x)\\ \cdot & & \cdot\\ \cdot & & \cdot\\ \cdot & & \cdot\\ \partial_1 f_m(x) & & \partial_n f_m(x)\end{pmatrix}$ *and* $h=\begin{pmatrix}h_1\\ \cdot\\ \cdot\\ \cdot\\ h_n\end{pmatrix}$ *then we can write using matrix multiplication $Df(x)(h)=\partial_{\{1,\ldots,n\}}(f_1,\ldots,f_m)(x)\cdot h$. The matrix $\frac{\partial(f_1,\ldots,f_m)(x)}{\partial(1,\ldots,n)}$ is called the Jacobian matrix of f at x.*

Proof. Using 14.35 we have $(Df(x)(h))_j = (\pi_j \circ Df(x))(h) = Df_j(x)(h)$. As f_j: $U \subseteq \mathbb{K}^n \rightarrow Y_j$ we have by 14.34 that $Df_j(x)(h) = \sum_{i \in \{1, \ldots, n\}} h_i \cdot \partial_i f_j(x)$ □

Example 14.37. Let $\langle \mathbb{K}^n, \| \|_n \rangle$, $\langle \mathbb{K}^m, \| \|_m \rangle$ be two vector spaces based on $\langle \mathbb{K}, \| \| \rangle$ (with the maximum norm), $\{e_i\}_{i \in \{1, \ldots, n\}}$, $\{f_i\}_{i \in \{1, \ldots, n\}}$ the canonical bases (see 10.161), , and $L: \mathbb{K}^n \rightarrow \mathbb{K}^m$ a linear mapping with matrix (see 10.300) $\mathcal{M}(L, \{e_i\}_{i \in \{1, \ldots, n\}}, \{f_i\}_{i \in \{1, \ldots, m\}}) = \{L_{i,j}\}_{(i,j) \in \{1 \ldots. m\} \times \{1, \ldots, n\}}$ in the canonical bassis . Then $\forall x \in \mathbb{K}^n$ L is differentiable at x with Jacobian $\frac{\partial(L_1, \ldots, L_m)(x)}{\partial(1, \ldots, n)} = \{L_{i,j}\}_{(i,j) \in \{1 \ldots. m\} \times \{1, \ldots, n\}}$ TOD check if we must not transpose

Proof. As L is linear it is differentiable at x (see 14.15) and $\forall j \in \{1, \ldots, m\}$ π_j is linear we have that $\pi_j \circ L$ is linear. Hence given $i \in \{1, \ldots, n\}, j \in \{1, \ldots, m\}$ we have that

$$\begin{array}{rcl} \partial_i L_j(x) & = & \partial_i(\pi_j \circ L)(x) \\ & \underset{14.34}{=} & (D(\pi_j \circ L)(x))(e_i) \\ & \underset{14.15}{=} & (\pi_j \circ L)(e_i) \\ & = & (L(e_i))_j \\ & \underset{10.300}{=} & L_{i,j} \end{array}$$

proving that $\frac{\partial(L_1, \ldots, L_m)(x)}{\partial(1, \ldots, n)} = \{L_{i,j}\}_{(i,j) \in \{1 \ldots. m\} \times \{1, \ldots, n\}}$. □

Theorem 14.38. *Let $\langle X_1, \| \|_1 \rangle$, $\langle X_2, \| \|_2 \rangle$, $\langle Y, \| \|_Y \rangle$ be normed spaces, and $\langle X_1 \times X_2, \| \| \rangle$ equipped with the maximum norm, $x = (x_1, x_2) \in X_1 \times X_2$ and $L \in L(X_1, X_2; Y)$ (a bilinear continuous mapping) then $L: X_1 \times X_2 \rightarrow Y$ is differentiable at x and $DL(x_1, x_2) = L(x_1, *) + L(*, x_2)$ where $L(x_1, *)(h_1, h_2) = L(x_1, h_2)$ and $L(*, x_2)(h_1, h_2) = L(h_1, x_2)$*

Proof. First we prove that $L(x_1, *) + L(*, x_2)$ is linear.

$$\begin{array}{rcl} (L(x_1, *) + L(*, x_2))(\alpha \cdot (h_1, h_2) + \beta \cdot (g_1, g_2)) & = & (L(x_1, *) + L(*, x_2))(\alpha \cdot h_1 + \beta \cdot g_1, \alpha \cdot h_2 + \beta \cdot g_2) \\ & = & L(x_1, \alpha \cdot h_2 + \beta \cdot g_2) + L(\alpha \cdot h_1 + \beta \cdot g_1, x_2] \\ & = & \alpha \cdot L(x_1, h_2) + \alpha \cdot L(h_1, x_2) + \beta \cdot L(x_1, g_2) + \beta \cdot L(g_1, x_2) \\ & = & \alpha \cdot (L(x_1, *) + L(*, x_2))(h_1, h_2) + \beta \cdot (L(x_1, *) + L(*, x_2))(g_1, g_2) \end{array}$$

Second we prove that $L(x_1, *) + L(*, x_2)$ is continuous let $(h_1, h_2) \in X_1 \times X_2$ with $\|(h_1, h_2)\| = 1$ then (see definition of maximum norm) we have $\|h_1\|_1 \leqslant 1$ and $\|h_2\|_2 \leqslant 1$, then $\|(L(x_1, *) + L(*, x_2))(h_1, h_2)\|_Y = \|L(x_1, h_2) + L(h_1, x_2)\|_Y \leqslant \|L(x_1, h_2)\|_Y + \|L(h_1, x_2)\|_Y \leqslant \|L\| \cdot \|x_1\|_1 \cdot \|h_2\|_2 + \|L\| \cdot \|x_2\|_2 \cdot \|h_1\|_1 \leqslant \|L\| \cdot (\|x\|_1 + \|x_2\|_2)$ which by 12.188 means that continuity is proved.

Finally let $\varepsilon > 0$ and take $0 < \delta < \frac{\varepsilon}{\|L\| + 1}$ then if $h = (h_1, h_2) \in (X_1 \times X_2)_x = X_1 \times X_2$ with $\|(h_1, h_2)\| < \delta \Rightarrow \|h_1\|_1, \|h_2\|_2 \leqslant \|(h_1, h_2)\| < \frac{\varepsilon}{\|L\| + 1}$ then

$$
\begin{aligned}
&\|L((x_1, x_2) + (h_1, h_2)) - L((x, y)) - \\
&(L(x_1, *) + L(*, x_2))(h_1, h_2)\|_Y &=& \ \|L(x_1 + h_1, x_2 + h) - L(x, y) - L(x_1, h_2) - L(h_1, x_2)\|_Y \\
&&=& \ \|L(x_1, x_2 + h_2) + L(h_1, x_2 + h_2) - L(x, y) - L(x_1, h_2) - L(h_1, x_2)\|_Y \\
&&=& \ \|L(x_1, x_2) + L(x_1, h_2) + L(h_1, x_2) + L(h_1, h_2) - L(x_1, x_2) - L(x_1, h_2) - L(h_1, x_2)\|_Y \\
&&=& \ \|L(h_1, h_2)\|_Y \\
&&\leqslant& \ \|L\| \cdot \|h_1\|_1 \cdot \|h_2\|_2 \\
&&<& \ \|L\| \cdot \frac{\varepsilon}{\|L\| + 1} \cdot \|h_2\|_2 \\
&&<& \ \varepsilon \cdot \|h_2\|_2 \\
&&\leqslant& \ \varepsilon \cdot \|(h_1, h_2)\|
\end{aligned}
$$

proving that $\mathrm{DL}(x, y) = L(x_1, *) + L(*, x_2)$ □

We can use the above theorem to proof the differential of a product of functions

Corollary 14.39. *Let $\langle X, \|\|_X \rangle$ be a normed space, $\langle \mathbb{K}, || \rangle$ ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$), $U \subseteq K$ open with $x \in U$ and $f, g: U \to \mathbb{K}$ two maps differentiable at x then $f \cdot g: U \to \mathbb{K}$ is differentiable at x with $D(f \cdot g)(x) = g(x) \cdot \mathrm{Df}(x) + f(x) \cdot \mathrm{Dg}(x)$.*

Proof. Define $h: X \to (\mathbb{K}, \mathbb{K})$ by $x \to (f(x), g(x))$ which is differentiable at $x \in U$ as $h_1 = \pi_1 \circ h = f$ and $h_2 = \pi_2 \circ h = g$ are differentiable (see 14.35). Also $p: \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ defined by $p(x, y) = x \cdot y$ is clearly bi-linear and continuous [as if $\|(x, y)\| \leqslant 1 \Rightarrow |x|$, $|y| \leqslant 1$ proving $|x \cdot y| \leqslant |x| \cdot |y| \leqslant 1$ and thus that p is continuous] so that it is differentiable with $Dp(x, y) = p(x, *) + p(*, y) = x \cdot * + * \cdot y$. Now as $f \cdot g = p \circ h$ we have by the chain rule (see 14.27) that $f \cdot g$ is differentiable and $D(f \cdot g)(x) = Dp(h(x)) \circ Dh(x) = Dp(f(x), g(x)) \circ Dh(x) = (f(x) \cdot * + * \cdot g(x)) \circ (Df(x), Dg(x))$ so that $D(f \cdot g)(x)(h) = (f(x) \cdot * + * \cdot g(x))(Df(x)(h), Dg(x)(h)) = f(x) \cdot Dg(x)(h) + g(x) \cdot Df(x)(h) \Rightarrow D(f \cdot g)(x) = f(x) \cdot Dg(x) + g(x) \cdot Df(x)$ □

If we take in the above theorem $\langle X, \|\|_X \rangle = \langle \mathbb{K}, || \rangle$ and 14.17 the above corollary gives

Corollary 14.40. *Take the normed space $\langle \mathbb{K}, || \rangle$, $U \subseteq \mathbb{K}$ a open set, $x \in U$ and f, $g: U \to \mathbb{K}$ functions that have a derivative at x then $f \cdot g$ has a derivative at x and $(f \cdot g)'(x) = f(x) \cdot g'(x) + g(x) \cdot f'(x)$*

Proof. Using 14.17 we have that f, g are differentiable and thus by the above theorem we have that $f \cdot g$ is differentiable and thus by 14.17 has a derivative. Also using 14.17 we have $(f \cdot g)'(x) = D(f \cdot g)(x)(1) = g(x) \cdot Df(x)(1) + f(x) \cdot Dg(x)(1) = g(x) \cdot f'(x) + f(x) \cdot g'(x)$ $\square$

Corollary 14.41. *Take the normed space $\langle \mathbb{K}, || \rangle$, $U \subseteq \mathbb{K}$ a open set, $x \in U$, $z \in \mathbb{N}$ define then $*^z: \mathbb{K} \to \mathbb{K}$ by $x \to x^z$ then $*^z$ has a derivative at $x \in \mathbb{K}$ given by $(*^z)'(x) = z \cdot *^{z-1}(x)$, if $z \in \{-z | z \in \mathbb{N}\}$ then $*^z: \mathbb{K} \setminus \{0\} \to \mathbb{K} \setminus \{0\}$ defined by $x \to x^z = \frac{1}{x^{-z}}$ has a derivative at $x \in \mathbb{K} \setminus \{0\}$ (a open set) given by $(*^z)'(x) = z \cdot *^{z-1}(x)$*

Proof. We have the following cases to consider:

1. **($z \in \mathbb{N}$)** We proceed here by induction so let $S = \{z \in \mathbb{N} | *^z \text{ has } a \text{ derivate } z \cdot *^{z-1}\}$ then we have

   $1 \in S$. If $z = 1$ then $*^1: \mathbb{K} \to \mathbb{K}$ is linear so that $(*^1)'(x) = D *^1(x)(1) = *^1(1) = 1^1 = 1 = 1 \cdot x^0 = 1 \cdot x^{z-1}$ so that $1 \in S$

   $z \in S \Rightarrow z + 1 \in S$. Assume now that $z \in S$ then we have that $*^{(z+1)}(x) = x^{z+1} = x \cdot x^z$ so that $*^{z+1} = *^1 \cdot *^z$ and then using the previous corollary we have $(*^{z+1})'(x) = (*^1 \cdot *^z)'(x) = x^1 \cdot (*^z)'(x) + (*^z)(x) \cdot (*^1)'(x) \underset{(a) \text{ and } z \in S}{=} x \cdot z \cdot (*^{z-1})(x) + x^z = x \cdot z \cdot x^{z-1} + x^z = z \cdot x^z + x^z = (z+1) \cdot x^z = (z+1) \cdot (*)^{(z+1)-1}(x)$ proving that $z + 1 \in S$

   using mathematical induction we have then $S = \mathbb{N}$

2. **($z \in \{-z | z \in \mathbb{N}\}$)** again we prove this by induction so let $S = \{z \in \mathbb{N} | *^{-z} \text{ has } a \text{ derivate } (-z) \cdot *^{(-z-1)}\}$ then we have

   $1 \in S$. If $z = 1$ then $*^{-1}: \mathbb{K} \setminus \{0\} \to \mathbb{K} \setminus \{0\}$ is defined by $x \to \frac{1}{x}$ now if $x \in \mathbb{K} \setminus \{0\}$ and h so that $x + h \in \mathbb{K} \setminus \{0\}$ then $\frac{\frac{1}{x+h} - \frac{1}{x}}{h} + \frac{1}{x^2} = \frac{\frac{x}{x \cdot (x+h)} - \frac{x+h}{x \cdot (x+h)}}{h} + \frac{1}{x^2} = \frac{\frac{-h}{x \cdot (x+h)}}{h} + \frac{1}{x^2} = \frac{-1}{x \cdot (x+h)} + \frac{1}{x^2} = \frac{1}{x} \cdot \left(\frac{-1}{x+h} + \frac{1}{x} \right) = \frac{1}{x} \cdot \frac{-x + x + h}{x \cdot (x+h)} = \frac{h}{x^2 \cdot (x+h)}$ giving

   $$\frac{\frac{1}{x+h} - \frac{1}{x}}{h} + \frac{1}{x^2} = \frac{h}{x^2 \cdot (x+h)} \tag{14.22}$$

   Given $\varepsilon > 0$ take then $0 < \delta < \min\left(\frac{\varepsilon \cdot |x|^3}{2}, \frac{|x|}{2} \right)$ then if $0 < |h| < \delta$ such that $x + h \in \mathbb{K} \setminus \{0\}$ (meaning $h \in (\mathbb{K} \setminus \{0\})_x$) we have $|h| < \frac{|x|}{2}$ giving $-\frac{|x|}{2} < -|h| \Rightarrow \frac{|x|}{2} = |x| - \frac{|x|}{2} < |x| - |h| \leqslant |x + h|$ [because $|x| = |x + h - h| \leqslant |x + h| + |h| \Rightarrow |x| - |h| \leqslant |x + h|$] and thus $\frac{|x|}{2} < |x + h|$ or $\frac{1}{|x+h|} < \frac{2}{|x|}$ and thus $\left| \frac{h}{x^2 \cdot (x+h)} \right| = \frac{|h|}{|x|^2 \cdot |x+h|} < \frac{2 \cdot |h|}{|x| \cdot |x|^2} = 2 \cdot \frac{|h|}{|x|^3} \leqslant 2 \cdot \frac{\varepsilon \cdot |x|^3}{|x^3| \cdot 2} = \varepsilon$ [as $|h| < \delta < \frac{\varepsilon \cdot |x|^3}{2}$]. So if $0 < \delta < \min\left(\frac{\varepsilon \cdot |x|^3}{2}, \frac{|x|}{2} \right)$ then $\left| \frac{*^{-1}(x+h) - *^{-1}(x)}{h} - (-1 \cdot *^{(-2)}(x)) \right| = \left| \frac{\frac{1}{x+h} - \frac{1}{x}}{h} - \frac{1}{x^2} \right| \underset{14.22}{=} \left| \frac{h}{x^2 \cdot (x+h)} \right| < \varepsilon$ proving that $(*^{-1})'(x) = -1 \cdot (*^{-1-1})(x)$ or $1 \in S$.

$\boldsymbol{z \in S \Rightarrow z+1 \in S}$**.** If $z \in S$ then $(*)^{-(z+1)}(x) = \frac{1}{*^{z+1}(x)} = \frac{1}{x^{z+1}} = \frac{1}{x \cdot x^z} = (*^{-1})(x) \cdot (*^{-z})(x)$ so that $(*)^{-(z+1)} = (*)^{-1} \cdot (*^{-z})$ so using the previous corollary we have $((*)^{-(z+1)})'(x) = (*^{-1})(x) \cdot (*^{-z})'(x) + (*^{-z})(x) \cdot (*^{-1})'(x) \underset{(a)\,\text{and}\,z \in S}{=} x^{-1} \cdot (-z) \cdot (*^{-z-1})(x) + x^{-z} \cdot (-1) \cdot (*^{-2})(x) = x^{-1} \cdot (-z) \cdot x^{-z-1} - x^{-z} \cdot x^{-2} = -z \cdot x^{-z-2} - x^{-z-2} = (-z-1) \cdot x^{-z-2} = (-(z+1)) * x^{-(z+1)-1} = (-(z+1)) \cdot (*^{-(z+1)-1})(x)$ proving that $(*^{-(z+1)})' = (-(z+1)) \cdot *^{-(z+1)-1}$ or $z+1 \in S$

Using mathematical induction we have then $S = \mathbb{N}$ □

14.2 Higher order differentiability

We define higher order differentiability using recursion as follows

Definition 14.42. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be normed spaces, $U \subseteq X$ a open set, $V \subseteq Y$ $n \in \mathbb{N}$ and $f: U \rightarrow V$ a function then f is n-times differentiable on U with n-the differential $D^n f: U \rightarrow L^n(X;Y)$ if one of the following is correct:*

1. *If $n = 1$ then $f: U \rightarrow V$ must be differentiable $\forall x \in U$ and $D^1 f: U \rightarrow L(X, Y)$ is defined by $x \rightarrow D^1 f(x) = Df(x)$*
2. *If $1 < n$ then f must be $(n-1)$-times differentiable on U and $D^{n-1} f: U \rightarrow L^{n-1}(X;Y)$ must be differentiable at every $x \in U$. $D^n f: U \rightarrow L^n(X;Y)$ is defined by $x \rightarrow D^n f(x) = D(D^{n-1} f)(x) \in L(X, L(X^{n-1};Y)) = L^n(X;Y)$ (see 12.199)*

Remark 14.43. We have seen by 12.213 that there exists a a norm preserving isomorphism $\mathcal{P}_n$: $L^n(X;Y) \rightarrow L(X^n;Y)$ allowing us to consider $\mathcal{P}_n(D^n f(x)) \in L(X^n;Y)$ as a multi-linear mapping where $\mathcal{P}_n(D^n f(x))(x_1, \ldots, x_n) = D^n f(x)(x_1: \ldots: x_n)$. To simplify notation when we write $D^n f(x)(h_1, \ldots, h_n)$ we actually means $\mathcal{P}_n(D^n f())(h_1, \ldots, h_n) = D^n f(x)(h_1: \ldots: h_n)$.

Definition 14.44. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be normed spaces, $U \subseteq X$ a open set, $V \subseteq Y$ $n \in \mathbb{N}$ and $f: U \rightarrow V$ a function then f is ∞-differentiableif $\forall n \in \mathbb{N}$ we have that f is n-differentiable.*

We next define the concept of differentiable classes

Definition 14.45. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be normed spaces, $U \subseteq X$ a open set, $V \subseteq Y$ $n \in \mathbb{N}_0$ and $f: U \rightarrow V$ a function then f is of class C^n if one of the following is satisfied*

1. *If $n = 0$ then f must be continuous on U*
2. *If $1 < n$ then f must be n-times differentiable on U and $D^n f: U \rightarrow L^n(X;Y)$ must be continuous on U*

Definition 14.46. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be normed spaces, $U \subseteq X$ a open set, $V \subseteq Y$ and $f: U \rightarrow V$ a function then f is of class C^∞ if $\forall n \in \mathbb{N}_0$ we have that f is of class C^n. Functions of class C^∞ are also called* ***smooth*** *functions.*

As for derivatives we use the following definition

Definition 14.47. *Let $\langle Y, \|\|_Y\rangle$, be a normed space and $\langle \mathbb{K}, \|\|\rangle$ be the normed space $\mathbb{K}$ equipped with the canonical norm, $U \subseteq \mathbb{K}$ a open set, $V \subseteq \mathbb{K}$, $n \in \mathbb{N}$ and $f: U \rightarrow V$ a function then f has a n-the derivative $f^{(n)}$ on U if one of the following is satisfied.*

1. *If $n = 1$ we must have $\forall x \in U$ that f has a derivative at x and we define $f^{(1)}: U \rightarrow Y$ by $x \rightarrow f^{(1)}(x) = f'(x)$*
2. *If $n > 1$ we must have that f has a $(n-1)$-the derivative on U and $\forall x \in U$ we have that $f^{(n-1)}: U \rightarrow Y$ has a derivative at x and we define $f^{(n)}: U \rightarrow Y$ by $x \rightarrow f^n(x) = (f^{(n-1)}(x))'$*

Lemma 14.48. *Let $\langle \mathbb{K}, \|\|\rangle$ be the vector space $\mathbb{K}$ using the canonical norm, $\langle Y, \|\|_Y\rangle$ a normed space, $U \subseteq \mathbb{K}$ a open set and $f: U \subseteq \mathbb{K} \rightarrow L(\mathbb{K}, Y)$ then if we define $f(*)(1): U \rightarrow L(\mathbb{K}, Y)$ by $t \rightarrow (f(*)(1))(t) = f(t)(1)$ then we have that f is continuous iff $f(\star)(1)$ is continous*

Proof.

1. **(⇒)** Let $x \in U$ then as f is continuous on U, f is continuous at x, so $\forall \varepsilon > 0$ there exists a $\delta > 0$ such that if $|x - y| < \delta \Rightarrow \|f(x) - f(y)\|_{L(\mathbb{K},Y)} < \varepsilon \Rightarrow \|(f(\star)(1))(x) - (f(*)(1))(y)\|_Y = \|f(x)(1) - f(y)(1)\|_Y = \|(f(x) - f(y))(1)\|_Y \leqslant \|f(x) - f(y)\|_{L(\mathbb{K},Y)} \cdot |1| = \|f(x) - f(y)\|_{L(\mathbb{K},Y)} < \varepsilon$ proving continuity of $f(*)(1)$ at x. As x was chosen arbitrary we have that $f(*)(1)$ is continuous on U.
2. **(⇐)** Let $x \in U$ then as $f(*)(1)$ is continuous on U we have that $f(*)(1)$ is continuous at x, so $\forall \varepsilon > 0$ there exists a $\delta > 0$ such that if $|x - y| < \delta \Rightarrow \|(f(*)(1))(x) - (f(*)(1))(y)\|_Y < \frac{\varepsilon}{2}$, then as $\|(f(x) - f(y))(r)\|_Y = \|(f(x) - f(y))(r \cdot 1)\|_Y = \|r \cdot (f(x) - f(y))(1)\|_Y = |r| \cdot \|(f(x) - f(y))(1)\|_Y = |r| \|f(x)(1) - f(y)(1)\|_Y = |r| \|(f(*)(1))(x) - (f(*)(1))(y)\|_Y < |r| \cdot \frac{\varepsilon}{2} \Rightarrow \|(f(x) - f(y))(r)\|_{L(\mathbb{K},Y)} \leqslant \frac{\varepsilon}{2}|r| < \varepsilon \cdot |r|$ proving that f is continuous at x. As x was chosen arbitrary we have that f is continuous on U. □

Theorem 14.49. *Let $\langle \mathbb{K}, \|\|\rangle$ be the vector space $\mathbb{K}$ using the canonical norm, $\langle Y, \|\|_Y\rangle$ a normed space, $U \subseteq \mathbb{K}$ a open set, $V \subseteq Y$ and $f: U \subseteq \mathbb{K} \rightarrow V$ a function then the following are equivalent:*

1. *f is C^1*
2. *$\forall t \in U$ we have that $f'(t) = D^1 f(x)(1)$ exists and $f': U \rightarrow Y$ defined by $t \rightarrow f'(t)$ is continuous on U [or f has a 1-derivative that is continuous).*

Proof.

1. **(1 ⇒ 2)** As f is C^1 we have that f is differentiable on U and $Df: x \rightarrow Df(x) \in L(\mathbb{K}, Y)$ is continuous. Then using 14.17 we have that $\forall x \in U$ $f'(x)$ exists and $f'(x) = Df(x)(1) = (Df(*)(1))(x)$ proving that $f' = Df(*)(1)$. We can use now the previous lemma to find that f' is continuous on U because continuity of Df.
2. **(2 ⇒ 1)** As $\forall x \in U$ $f'(x)$ exists we use 14.17 again to find the existence of $Df(x)$ where $f'(x) = Df(x)(1)$ and using the previous lemma and the continuity of $f' = Df(*)(1)$ we find that $Df: x \rightarrow Df(x)$ is continuous. □

Theorem 14.50. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces, $U \subseteq X$ a open set, $V \subseteq Y$ and $f: U \rightarrow V$ a function then f is of class C^∞ iff f is ∞-times differentiable.*

Proof.

1. **($\Rightarrow$)** If $\forall n \in \mathbb{N}_0$ we have that f is of class C^n then by definition we have that f is n-times differentiable proving that f is ∞-times differentiable.
2. **($\Leftarrow$)** If $n = 0$ then as f is differentiable at every $x \in U$ we have by 14.10 that f is continuous at x and thus that f is continuous and thus C^0. If $n \in \mathbb{N}$ we have that f is n-times differentiable then f is $(n+1)$-times differentiable which as $\forall x \in U$ we have $D^{n+1} f(x) = D(D^n f)(x)$ proves by 14.10 that $D^n f$ is continuous at x and thus that $D^n f$ is continuous on U proving that f is of class C^n. So we have that f is of class C^∞. □

Theorem 14.51. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces, $U \subseteq X$ a open set, $V \subseteq Y$, $n \in \mathbb{N}_0$ and $f: U \rightarrow V$ a function that is n-times differentiable then we have that $\forall m \in \{1, \ldots, n\}$ f is m-times differentiable*

Proof. We proceed by induction so let $S_n = \{i \in \mathbb{N}_0 | \text{if } 0 \leqslant i < n \text{ then } f \text{ is } (n-i)\text{-times differentiable}\}$ we have then

$0 \in S_n$. f $i = 0 \Rightarrow n - i = n$ and as by the hypothesis f is n-times differentiable we have $0 \in S_n$

$i \in S \Rightarrow i + 1 \in S$. If $i \in S_n$ and $0 \leqslant i + 1 < n$ (so that $1 < (n - i)$) then as $0 \leqslant i < n$ and $n \in S$ we have that f is $(n-i)$-times differentiable on U, so by definition (as $1 < (n-1)$) we have that f is $((n-i)-1)$-times differentiable and as $((n-i)-1) = (n-(i+1))$ we have that $i + 1 \in S_n$

By mathematical induction we have then that $S_n = \mathbb{N}_0$. So if $m \in \{1, \ldots, n\}$ then $i = n - m \in \{0, \ldots, n-1\}$ proving as $i \in \mathbb{N}_0 = S_n$ and $0 \leqslant i < n$ that f is $(n-i)$-times differentiable which proves our theorem as $m = n - i$. □

Theorem 14.52. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces, $U \subseteq X$ a open set, $V \subseteq Y$, $n \in \mathbb{N}_0$ and $f: U \rightarrow V$ a function of class C^n then $\forall i \in \{0, \ldots, n\}$ we have that f is of class C^i*

Proof. We have to consider the following cases for n:

$n = 0$. Then $\{0, \ldots, n\} = \{0\}$ which makes the theorem trivial

$n = 1$. Then f is of class C^1 and thus f is differentiable everywhere on U so that by 14.10 f is continuous on U and thus C^0.

$1 < n$. Two cases must be considered:

$m = n$. then by the hypothesis we have that f is of class C^m

$m < n$. Then by the previous theorem f is m-times and $(m+1)$-times differentiable on U, so that $D^m f: U \rightarrow L^m(X; Y)$ is differentiable at every $x \in U$, the $(m+1)$-times differentiability together with 14.10 means that $D^m f$ is continuous on U and thus that f is C^m □

The next theorem is very usable in induction proofs.

Theorem 14.53. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces, $U \subseteq X$ a open set, $V \subseteq Y$, $n \in \mathbb{N} \setminus \{1\}$ and $f: U \to V$ a function then f is n-times differentiable if and only if $D^1 f$ is $(n-1)$-times differentiable, also $D^{n-1}(D^1 f) = D^n f$*

Proof.

⇒. This is proved by induction so let $S = \{n \in \{2, \ldots\} | \text{if } a \text{ function } h: U \to V \text{ is } n-$ times differentiable then Dh is $(n-1)$ times differentiable and $D^{n-1}(Dh) = D^n h\}$ then we have:

$2 \in S$. If $n = 2$ then as f is 2-times differentiable we have that f is 1-times differentiable and $D^1 f: U \to L(X, Y)$ is differentiable at every $x \in U$ proving that $D^1 f$ is 1-times differentiable, also $D^2 f: U \to L^2(X; Y)$ is defined by $D^2 f(x) = D(D^1 f)(x) = D^1(D^1 f)(x) \Rightarrow D^2 f = D^1(D^1 f)$ so $2 \in S$

$n \in S \Rightarrow n + 1 \in S$. If $n \in S$ and let f be $(n+1)$-times differentiable then f is by definition n-times differentiable on U, so as $n \in S$ we have that $D^1 f$ is $(n-1)$-times differentiable and $D^n f = D^{n-1}(D^1 f)$. From the $(n+1)$-differentiability of f we have $\forall x \in U$ that $D^n f$ is differentiable at x and $D^{n+1} f(x) = D(D^n f)(x) = D(D^{n-1}(D^1 f)(x)) = D^n(D^1 f)(x) \Rightarrow D^{n+1} f = D^n(D^1 f)$ proving that $n + 1 \in S$

By induction we have then $S = \mathbb{N} \setminus \{1\}$ porving the theorem.

⇐. Again we use induction so let $S = \{n \in \{2, \ldots\} |$ if a function $h: U \to V$ is such that $D^1 h$ is $(n-1)$ -times differentiable then h is n-times differentiable and $D^{n-1}(D^1 h) = D^n h\}$ then:

$2 \in S$. If n=2 then as $D^1 f$ is $2 - 1 = 1$-times differentiable we have that $\forall x \in U$ we have that $D^{\,1} f(x)$ exists and $D^1 f: U \to L(X, Y)$ is differentiable at every $x \in U$ so that by definition f is 2 times differentiable and $D^2 f: U \to L(X^2; Y)$ is defined by $x \to D^2 f(x) = D(D^1 f)(x)$ proving that $D^2 f = D^1(D^1 f)$ and thus that $2 \in S$

$n \in S \Rightarrow n + 1 \in S$. Let $n \in S$ then if $D^1 f$ is $(n+1) - 1 = n$ times differentiable, by definition we have then that $D^1 f$ is $(n-1)$-times differentiable and that $D^{n-1}(D^1 f): U \to L^{n-1}(X; L(X, Y)) \underset{12.202}{=} L^n(X; Y)$ defined by $x \to D^{n-1}(D^1 f)(x)$ is differentiable everywhere on U and $D^n(D^1 f): U \to L(X, L^n(X; Y)) = L^{n+1}(X; Y)$ is defined by $x \to D^n(D^1 f)(x) = D(D^{n-1}(D^1 f))(x)$ proving that

$$D^{n-1}(D^1 f) \text{ is differentiable on } U \tag{14.23}$$

$$D^n(D^1 f) = D^1(D^{n-1}(D^1 f)) \tag{14.24}$$

As $n \in S$ we have that f is n-times differentiable with $D^n f = D^{n-1}(D^1 f)$ which by 14.23 means that $D^n f$ is differentiable, proving that f is $(n+1)$-times differentiable and $D^{n+1} f: U \to L^{n+1}(X; Y)$ is defined by $x \to D^{n+1} f(x) = D(D^n f)(x) = D(D^{n-1}(D^1 f))(x) = D^1(D^{n-1}(D^1 f))(x) \underset{14.24}{=} D^n(D^1 f)(x)$ so that $D^{n+1} f = D^n(D^1 f)$. This proves that $n + 1 \in S$.

Induction prove then that $S = \mathbb{N} \setminus \{1\}$ □

Next we extend the above to show that it also works for differentiable classes

Theorem 14.54. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces, $U \subseteq X$ a open set, $V \subseteq Y$ a open set, $n \in \mathbb{N}$ and $f: U \rightarrow Y$ a function then f is of class C^n if and only if $D^1 f$ exists and is of class C^{n-1}*

Proof. The following cases must be considered

$n = 1$. If f is of class C^1 then by definition $D f$ is continuous and thus of class C^0. If $D^1 f$ exists and is of class C^0 then f is 1-times differentiable with $D^1 f$ continuous and thus of class C^1.

$n > 1$.

- **$\Rightarrow$.** If f is of class C^n then f is n-times differentiable and $D^n f$ is continuous, using the previous theorem we have then that $D^1 f$ is $(n-1)$ times differentiable and $D^{n-1}(D^1 f) = D^n f$ which is continuous so that D^1 is of class C^{n-1}
- **$\Leftarrow$.** If $D^1 f$ is of class C^{n-1} then it is $n-1$ times differentiable and continuous so again by the previous theorem we have then that f is n-times differentiable and as $D^n f = D^{n-1}(D^1 f)$ which is continuous we have that f is of class C^n. □

In the next theorem we prove that n-times differentiability is a local property.

Theorem 14.55. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces, $U \subseteq X$ a open set, $W \subseteq Y$, $n \in \mathbb{N}$ and $f: U \rightarrow W$ a function then*

1. *If f is n-times differentiable on U then $\forall V \subseteq U$, V open we have $f_{|V}$ is n-times differentiable on V and $(D^n f)_{|V} = D^n(f_{|V})$*
2. *If $\forall x \in U$ there exists a open V with $x \subseteq V \subseteq U$ for which $f_{|V}: V \rightarrow W$ is n-times differentiable on V then f is n-times differentiable on U and $(D^n f)_{|V} = D^n(f_{|V})$*

Proof.

1. We proceed by induction, so let $S = \{i \in \mathbb{N} | \text{if } i \leqslant n | \text{if } V \subseteq U, V \text{ open then } f_{|V}: V \rightarrow W \text{ is } n-\text{times differentiable on } V \text{ and } D^i(f_{|V}) = (D^i f)_{|V}\}$ then we have

   $1 \in S$. If $i = 1$ then as $f: U \rightarrow W$ is 1-times differentiable we have by definition that $\forall x \in U$ we have that $f: U \rightarrow W$ is differentiable and $D^1 f$ is defined by $D^1 f(x) = D f(x)$. Using 14.12 we have then that $\forall x \in V$ that $f_{|V}: U \rightarrow W$ is differentiable at x with $D(f_{|V})(x) = D f(x)$ or in other words $f_{|V}: U \rightarrow W$ is 1-times differentiable and $D^1(f_{|V})$: $V \rightarrow L(X, Y)$ is defined by $D^1(f_{|V})(x) = D(f_{|V})(x) = D f(x) = D^1 f(x)$ or $D^1(f_{|V}) = (D^1 f)_{|V}$. So $1 \in S$

   $i \in S \Rightarrow i + 1 \in S$. If $i \in S$ then if $i + 1 \leqslant n$ we have that $D^i f: U \rightarrow L(X^i; Y)$ is differentiable $\forall x \in U$. So using 14.12 we have that $\forall x \in V$ that $(D^i f)_{|V}$ is differentiable at x which, as from $i \in S$ we have that $f_{|V}$ is i-times differentiable, means that $f_{|V}$ is $i + 1$-times differentiable on V. Also $\forall x \in V$ we have that $(D^{i+1}(f_{|V}))(x) = D(D^i(f_{|V}))(x) \underset{i \in S}{=} D((D^i f)_{|V})(x) \underset{14.12}{=} (D(D^i f))_{|V}(x) = (D^{i+1} f)_{|V}(x)$ proving that $D^{i+1}(f_{|V}) = (D^{i+1} f)_{|V}$. So we $i + 1 \in S$

Using mathematical induction we have then $S = \mathbb{N}$ proving (1)

2. We prove this by induction, so let $S = \{n \in \mathbb{N} | \text{if} \, \forall x \in u \vDash \exists V$ open with $x \in V \subseteq U$ such that $f_{|V}$ is $n -$ times differentiable then f is $n -$ times differentiable and $(D^n f)_{|V} = D^n(f_{|V})\}$, we have then

   $1 \in S$. By assumption $\forall x \in U$ there exists a V open with $x \in V \subseteq U$ such that $f_{|V}$ is differentiable at x, using 14.12 we have then that also f is differentiable at x and $Df(x) = Df_{|V}(x) \; \forall x \in V \Rightarrow (D^1 f)_{|V} = D^1(f_{|V})$ proving that $1 \in V$

   $n \in S \Rightarrow n+1 \in S$. Let $n \in S$, assume now that $\forall x \in U$ there exists a V open with $x \in V \subseteq U$ and $f_{|V}$ be $(n+1)$-times differentiable. By definition we have then that $f_{|V}$ is n-times differentiable and thus as $n \in S$ we have that f is n-times differentiable with $(D^n f)_{|V} = D^n(f_{|V})$. Also by definition and the fact that $f_{|V}$ is $(n+1)$-times differentiable we have that $D^n(f_{|V}) = (D^n f)_{|V}$ is differentiable at every $y \in V$ (and also using 14.12 $D^n f$ is differentiable at every $y \in U$) with $D^{n+1}(f_{|V})(y) = D(D^n(f_{|V}))(y) = D((D^n f)_{|V})(y) \underset{14.12}{=} (D(D^n f))_{|V}(y) = (D^{n+1} f)_{|V}(y)$ proving that $(D^{n+1} f)_{|V} = D^{n+1}(f_{|V})$ and as $x \in U$ is choosen arbitrary also that f is $(n+1)$-times differentiable at $x \in U$. So we conclude that $n+1 \in S$

   induction proves then that $S = \mathbb{N}$ finishing the proof. □

Next we prove that the above theorem is also true for classes

Theorem 14.56. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be normed spaces, $U \subseteq X$ a open set, $n \in \mathbb{N}_0$, $W \subset Y$ and $f: U \rightarrow W$ a function then:*

1. *If $f: U \rightarrow W$ is n-times differentiable on U then if V is open with $V \subseteq U$ then $f_{|V}$ is of class C^r.*
2. *If $\forall x \in U$ there exists a V open with $x \in V \subseteq U$ for which $f_{|V}: V \rightarrow W$ is of class C^n then f is of class C^n.*

Further if $n > 0$ then $D^n(f_{|V}) = (D^n f)_{|V}$

Proof. We have to consider the following cases

$n = 0$.

1. This follows from 12.133.
2. This follows from 12.139.

$n > 0$.

1. Using the previous theorem we have that if V is open and $V \subseteq U$ then $f_{|V}$ is n-times differentiable with $D^n(f_{|V}) = (D^n f)_{|V}$ which is continous on V by 12.133.

2. By the hypothesis we have that $\forall x\in U$ there exists a V open with $x\in V\subseteq U$ for which $f_{|V}: V\rightarrow Y$ is of class C^n, as $f_{|V}$ is then by definition n-times differentiable we have by the previous theorem proved that f is n-times differentiable on U with $(D^n f)_{|V}=D^n(f_{|V})$. As by definition of C^n $D^n(f_{|V})=(D^n f)_{|V}$ is continuous we have by 12.139 that $D^n f$ is continuous which at last proves that f is of class C^n. □

Theorem 14.57. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces, $U\subseteq X$, $W_f,W_g\subseteq Y$ open sets, $n\in\mathbb{N}$ and $f: U\rightarrow W_f$, $g: U\rightarrow W_f$ n-times differentiable functions then $f+g: U\rightarrow W_f+W_g$ is n-times differentiable and $D^n(f+g)=D^n f+D^n g$. Also if $\alpha\in\mathbb{K}$ then $\alpha\cdot f: U\rightarrow\alpha\cdot W_f$ is n-times differentiable and $D^n(\alpha\cdot f)=\alpha\cdot D^n f$. Note that by 12.76 we have that $\alpha\cdot W_f$ and W_f+W_g are open sets.*

Proof. This is proved by induction so let $S=\{n\in\mathbb{N}|\text{if } f,g: U\rightarrow Y \text{ are } n-\text{times differentiable then } f+g,\alpha\cdot f \text{ is } n-\text{times differentiable with } D^n(f+g)=D^n f+D^n g, D^n(\alpha\cdot f)=\alpha\cdot D^n f\}$ then

1. If $n=1$ then by 14.26 we have that $1\in S$
2. Assume that $n\in S$ and that f,g are $(n+1)$ times differentiable functions. Then f,g are n-times differentiable and thus $f+g,\alpha\cdot f$ are n-times differentiable with $D^n(f+g)=D^n f+D^n g$, $D^n(\alpha\cdot f)=\alpha\cdot D^n f$. As by definition we have that $D^n f$, $D^n g$ are differentiable on U we can use 14.26 again to proof that $D^n f+D^n g=D^n(f+g)$ and $\alpha\cdot D^n f=D^n(\alpha\cdot f)$ are differentiable on U with $D^{n+1}(f+g)(x)=D(D^n(f+g))(x)=D(D^n f+D^n g)(x)=D(D^n f)(x)+D(D^n f)(x)=D^{n+1}f(x)+D^{n+1}g(x)$ and $D^{n+1}(\alpha\cdot f)=D(D^n(\alpha\cdot f))(x)=D(\alpha\cdot D^n f)(x)=\alpha\cdot D(D^n f)(x)=\alpha\cdot D^{n+1}f(x)$ proving that $(n+1)\in S$ □

Theorem 14.58. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces, $U\subseteq X$ a open set, $W_f,W_g\subseteq Y$ open sets. $n\in\mathbb{N}_0$ and $f: U\rightarrow W_f$, $g: U\rightarrow W_g$ are of class C^n then $f+g$: $U\rightarrow W_f+W_g$ and $\alpha\cdot f: U\rightarrow\alpha\cdot U_f$ are of class C^n.*

Proof. The following cases must be considered

1. $\boldsymbol{(n=0)}$ First using 12.134 and the fact that for $f: U\rightarrow Y$ and $g: U\rightarrow Y$ we have $f(U)\subseteq W_f$, $g(U)\subseteq W_g$ we have $f: U\rightarrow Y, g: U\rightarrow Y$ are continuous. As the the sum of continuous functions is continuous and the product of a continuous function with a scalar is continuous (see 12.170) we have that $f+g: U\rightarrow Y$ and $\alpha\cdot f: U\rightarrow Y$ is continuous. Using 12.134 again and the fact that $(f+g)(U)\subseteq W_f+W_g, \alpha\cdot f\subseteq\alpha\cdot U_f$ we have that $f+g: U\rightarrow W_f+W_g$ and $\alpha\cdot f: U\rightarrow\alpha\cdot W_f$ are continuous.
2. $\boldsymbol{(n>0)}$ Using the previous theorem we have already that $f+g$ and $\alpha\cdot f$ are n-times differentiable. As $D^n f$, $D^n g$ are also continuous we can use (1) to prove that $D^n(f+g)=D^n f+D^n g$ and $D^n(\alpha\cdot f)=\alpha\cdot D^n f$ are continuous proving that $f+g,\alpha\cdot f$ are of class C^n. □

Using mathematical induction it is easy to extend the above theorems to a finite sum

Theorem 14.59. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces $U \subseteq X$ a open set, $k \in \mathbb{N}$, $n \in \mathbb{N}_0$ $\{f_i: U \to W_i\}_{i \in \{1,\ldots,k\}}$ a finite set of functions from which are n-times differentiable $(n > 0)$ [or C^n] then $\sum_{i \in \{1,\ldots,k\}} f_i: U \to \sum_{i \in \{1,\ldots,k\}} W_i$ is n-times differentiable [or C^n]. Here given sets $\{W_i\}_{i \in \{1,\ldots,k\}}$ we recursively define $\sum_{i \in \{1,\ldots,k\}} W_i = \left\{ \begin{array}{l} W_1 \text{ if } k = 1 \\ (\sum_{i \in \{1,\ldots,k-1\}} W_i) + W_k \text{ if } k > 1 \end{array} \right.$.*

Proof. Let $S = \{k \in \mathbb{N} | \text{if } \{f_i: U \to W_i\}_{i \in \{1,\ldots,n\}}$ is *a* finite set of functions from U to Y which are $n-$ times differentiable $(n > 0)$ [or C^n] then $\sum_{i \in \{1,\ldots,k\}} f_i$: $\sum_{i \in \{1,\ldots,k\}} W_i$ is $n-$ times differentiable [or C^n]$\}$ then we have:

1. If $k = 1$ then $\sum_{i \in \{1,\ldots,1\}} f_i = f_1: U \to W_1 = \sum_{i \in \{1,\ldots,1\}} W_1$ which is 1-times differentiable [or C^1] so that $1 \in S$
2. If $k \in S$ then if $\{f_i: U \to W_i\}_{i \in \{1,\ldots,k+1\}}$ are n-times differentiable functions [or C^n] then we have $\sum_{i=1}^{k+1} f_i = (\sum_{i=1}^{k} f_i) + f_n: U \to \left(\sum_{i \in \{1,\ldots,k\}} W_i\right) + W_{k+1}$ which is n-times differentiale (by 14.57) [or C^n by 14.58] proving that $k + 1 \in S$ □

Theorem 14.60. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces $U \subseteq X$ open, $y \in Y$ then $C_{U,y}: U \to Y$ defined by $x \to C_{U,y}(x) = y$ then $C_{U,y}$ is C^∞ and $D^n C_{U,y} = C_{U,0_{L^{n-1}(X,Y)}}$ where$\left[0_{L^n(X,Y)} = \left\{ \begin{array}{l} \text{neutral element in } Y \text{ if } n = 0 \\ \text{neutral element in } L^n(X;Y) \text{ if } n > 0 \end{array} \right. \right]$.*

Proof. Let $S = \left\{ n \in \mathbb{N} | D^n C_{U,y} \text{ exists and } D^n C_{U,y} = C_{X,0_{L^{n-1}(X;Y)}} \right\}$ then we have

1. If $n = 1$ then $\forall x \in U$ we have $D^1 C_{U,y}(x) = D(C_{U,y}) \underset{14.14}{=} C_{X,0} = C_{X,0_{L^{1-1}(X;Y)}}$ so that $1 \in S$
2. If $n \in S$ then $\forall x \in U$ we have $D^n C_{U,y} = C_{X,0_{L^{n-1}(X;Y)}}$ so by 14.14 we have that $D^n C_{U,y}$ is differentiable at every $x \in U$ with $D^{n+1} C_{U,y}(x) = D(D^n C_{U,y})(x) = D\left(C_{X,0_{L^{n-1}(X;Y)}}\right)(x) = 0_{L^{(n+1)-1}(X;U)}$ so that $D^{n+1} C_{U,y} = C_{U,0_{L^{(n+1)-1}(X;Y)}}$ proving that $n + 1 \in S$.

By induction we have $S = \mathbb{N}$ proving our theorem. □

Theorem 14.61. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces over $\mathbb{K}$ then every linear continuous map $L \in L(X, Y)$ is C^∞ where $\forall x \in D^1 L(x) = L$ and if $n \in \mathbb{N}$, $n > 1$ then $D^n L = C_{X,0_{L^{n-1}(X;Y)}} \in L^n(X;Y)$*

Proof. Using 14.50 we only have to prove ∞-times differentiability so let $n \in \mathbb{N}$ then we have the following cases to consider:

1. **($n = 1$)** Then $\forall x \in X$ we have by 14.15 that $DL(x) = L$ exists so that $D^1 L = C_{X,L}$
2. **($n > 1$)** Using the above theorem and (1) we have that $D^1 L$ is ∞-times differentiable, so if $n > 1$ then $n - 1 > 0$ and $D^{n-1}(D^1 L) = C_{X,0_{L^{(n-1)-1}(X;L(X,Y))}} \underset{12.202 \Rightarrow L^{n-2}(X;L(X,Y)) = L^{n-1}(X;Y)}{=} C_{X,0_{L^{n-1}(X;Y)}}$. As by 14.53 we have $D^n L = D^{n-1}(D^1 L) = C_{X,0_{L^{n-1}(X;Y)}}$ □

Theorem 14.62. *Let $\langle X, \|\|_X\rangle$ be a normed space and $\tau_{x_0}: X \to X$ defined by $x \to \tau_{x_0}(x) = x + x_0$ then τ_{x_0} is C^∞ with $D^1\tau_{x_0} = C_{X,1_X}$ and if $n > 1$ then $D^n\tau_{x_0} = C_{X,0_{L^{n-1}(X;Y)}}$*

Proof. We must look at the following cases

1. **($n=1$)** then if $x \in X$ we have that $\forall \varepsilon > 0$ we have if $\delta = 1$ and $\|h\|_X < 1$ with $h \in X_x = X$ that $\|\tau_{x_0}(x+h) - \tau_{x_0}(x) - 1_X(h)\|_X = \|x_0 + x + h - x_0 - x - h\|_X = \|0\|_X = 0$ proving that $D\tau_{x_0}(x) = 1_X$ so that $D^1\tau_{x_0} = C_{X,1_X}$
2. **($n > 1$)** Using the theorem about constant functions (see 14.60) we have that $D^1\tau_{x_0}$ is ∞-times differentiable, so if $n > 1$ then $n - 1 > 0$ and $D^{n-1}(D^1\tau_{x_0}) = C_{X,0_{L^{(n-1)-1}(X;L(X,Y))}} \underset{12.202 \Rightarrow L^{n-2}(X;L(X,Y)) = L^{n-1}(X;Y)}{=} C_{X,0_{L^{n-1}(X;Y)}}$. As by 14.53 we have $D^n\tau_{x_0} = D^{n-1}(D^1\tau_{x_0}) = C_{X,0_{L^{n-1}(X;Y)}}$ □

Theorem 14.63. *Let $\langle X_1, \|\|_1\rangle, \langle X_2, \|\|_2\rangle$ and $\langle Y, \|\|_Y\rangle$ be normed spaces. $L \in L(X_1, X_2; Y)$ a continuous bilinear mapping from $X_1 \times X_2 \to Y$ then L is C^∞. Further we have*

1. *D^1L is defined by $(x, y) \to L(X_1 \times X_2, Y)$ by $D^1L(x, y) = L(x, *) + L(*, y)$ where $(L(x, *) + L(*, y)) \in L(X_1 \times X_2, Y)$ is defined by $(r, s) \to (L(x, *) + L(*, y))(r, s) = L(x, s) + L(r, y)$*
2. *$D^2L = C_{X_1 \times X_2, D^1L}$*
3. *If $n > 2$ then $D^nL = C_{X_1, X_2, 0_{L^{n-1}(X_1 \times X_2;Y)}}$*

Proof. We have to consider the following cases:

1. **($n=1$)** Using the differential of a bi linear mapping (see 14.38) we have that L is differentiable at every $(x, y) \in X_1 \times X_2$ with $D^1L(x, y) = \mathrm{DL}(x, y) = L(x, *) + L(*, y)$ so that D^1L is defined by $(x, y) - D^1L(x, y) = L(x, *) + L(*, y)$
2. **($n=2$)** We prove now that $D^1L: X_1 \times X_2 \to Y$ is linear and continuous.
   a. **(linearity)** If $\alpha, \beta \in \mathbb{K}$ and $(r, s), (x_1, y_1), (x_2, y_2) \in X_1 \times X_2$ then $D^1L((\alpha \cdot (x_1, y_1) + \beta \cdot (x_2, y_2)))(r, s) = D^1L(\alpha \cdot x_1 + \beta \cdot x_2, \alpha \cdot y_1 + \beta \cdot y_2)(r, s) = (L(\alpha \cdot x_1 + \beta \cdot x_2, *) + L(*, \alpha \cdot y_1 + \beta \cdot y_2))(r, s) = L(\alpha \cdot x_1 + \beta \cdot x_2, s) + L(r, \alpha \cdot y_1 + \beta \cdot y_2) \underset{L \,\mathrm{is\,multilinear}}{=} \alpha \cdot L(x_1, s) + \alpha \cdot L(r, y_1) + \beta \cdot L(x_2, s) + \beta \cdot L(r, y_2) = \alpha \cdot (L(x_1, *) + L(*, y_1))(r, s) + \beta \cdot (L(x_2, *) + L(*, y_2))(r, s) = \alpha \cdot D^1L(x_1, y_1)(r, s) + \beta \cdot D^1L(x_2, y_2)(r, s)$ proving that $D^1L((\alpha \cdot (x_1, y_1) + \beta \cdot (x_2, y_2))) = \alpha \cdot D^1L(x_1, y_1) + \beta \cdot D^1L(x^2, y_2)$ proving linearity.
   b. **(continuity)** If $(x, y)(r, s) \in X_1 \times X_2$ then we have $\|D_1L(x, y)(r, s)\|_Y = \|L(x, s) + L(r, y)\|_Y \leqslant \|L(x, s)\|_Y + \|L(r, y)\|_Y \leqslant \|L\| \cdot (\|x\|_{X_1} \cdot \|s\|_{X_2}) + \|L\| \cdot (\|y\|_{X_2} \cdot \|r\|_{X_1}) = \|L\| \cdot (\|x\|_{X_1} \cdot \|r\|_{X_1} + \|y\|_{X_2} \cdot \|s\|_{X_2}) \leqslant \|L\| \cdot ((\|x\|_{X_1} + \|y\|_{X_2}) \cdot \max\{(\|r\|_{X_1}), \|s\|_{Y_2}\}) = (\|L\| \cdot (\|x\|_X + \|y\|_Y)) \cdot \|(r, s)\|_{X_1 \times X_2}$ proving that $\|D_1L(x, y)\|_{L(X_1 \times X_2, Y)} \leqslant \|L\| \cdot (\|x\|_{X_1} + \|y\|_{X_2}) \leqslant \|L\| \cdot \max(\{\|x\|_X, \|y\|_{X_2}\}) = \|L\| \cdot \|(x, y)\|_{X_1 \times X_2}$ proving that D_1L is continuous.

Using then 14.61 we have that $D^1 L$ is differentiable and $D^2 L = D^1(D^1 L) = C_{X_1 \times X_2, D_1 L}$

3. **($n > 2$)** As we have just proved that $D^1 L$ is a continuous linear function and thus ∞-times differentiable (see 14.61) we have that $D^n L \underset{n-1>1 \text{ and } 14.53}{=} D^{n-1}(D^1 f) = D^{n-1}(C_{X_1 \times X_2, D_1 L}) \underset{14.61, n-1>1}{=} C_{X_1 \times X_2, 0_{L^{n-2}(X_1 \times X_2; L(X_1 \times X_2; Y))}} = C_{X_1 \times X_2, 0_{L^{n-1}(X_1 \times X_2; Y)}}$ □

Theorem 14.64. *Let $\{\langle X_i, \|\|_i \rangle\}_{i \in \{1, \ldots, n\}}$ be a finite family of normal spaces then the projection maps $\pi_i: \prod_{j \in \{1, \ldots, n\}} X_j \rightarrow X_i$ are C^∞*

Proof. This is trivial as the projection maps are linear and continuous (see 12.144) so that we can use 14.61. □

Theorem 14.65. *Let $\langle X, \|\|_X \rangle$ be a normed space, $\{\langle Y_i, \|\|_i \rangle\}_{i \in \{1, \ldots, n\}}$ a family of normed spaces and $\|\|$ the maximum norm defined on $\prod_{i \in \{1, \ldots, n\}} Y_i$ then given $U \subseteq X$, $V \subseteq Y = \prod_{i \in \{1, \ldots, n\}} Y_i$ open sets and $r \in \mathbb{N}$ we have that $f: U \rightarrow V$ is r-times differentiable if and only if $\forall i \in \{1, \ldots, n\}$ we have $\pi_i \circ f: U \rightarrow \pi_i(V)$ is r-times differentiable and then $\pi_i \circ D^n f = D^n(\pi_i \circ f)$*

Proof. We prove this by induction so let $S = \left\{r \in \mathbb{N} | \text{ if } f: U \rightarrow \prod_{i \in \{1, \ldots, n\}} Y_i \text{ then } f \text{ is } r-\text{times diffrentiable iff } \forall i \in \{1, \ldots, n\}\ \pi_i \circ f \text{ is } r-\text{times differentiable and then } \pi_i \circ D^n f = D^r(\pi_i \circ f)\right\}$ then:

1. If $r = 1$ then using 14.35 we find that $\forall x \in U$ that f is differentiable at $x \in U$ iff $\forall i \in \{1, \ldots, n\}$ $\pi_i \circ f$ is differentiable at x and that $\pi_i \circ \mathrm{Df}(x) = D(\pi_i \circ f)(x)$ so that $1 \in S$

2. If $r \in S$ consider then $r + 1 > 1$ then

   a. **($\Rightarrow$)** If f is $r + 1$ times differentiable then $D^1 f$ is n-times differentiable so that as $n \in S$ we have that $\pi_i \circ D^1 f$ is r-times differentiable with $D^r(\pi_i \circ D^1 f) = \pi_i \circ D^r(D^1 f)$. Using (1) we have then that $D^1(\pi_i \circ f) \underset{(1)}{=} \pi_i \circ D^1 f$ is r-times differentiable or using 14.53 we have that $\pi_i \circ f$ is $(r + 1)$-times differentiable with $\pi_i \circ D^{r+1} f = \pi_i \circ D^r(D^1 f) \underset{r \in S}{=} D^r(\pi_i \circ D^1 f) \underset{(1)}{=} D^r(D^1(\pi_i \circ f)) = D^{r+1}(\pi_i \circ f)$.

   b. **($\Leftarrow$)** Assume that $\pi_i \circ f$ is $(r + 1)$-times differentiable and $\pi_i \circ D^{r+1} f = D^{r+1}(\pi_i \circ f)$, then $D^1(\pi_i \circ f)$ is r-times differentiable by 14.53 and thus as $r \in S$ we have that $D^1 f$ is r-times differentiable, using 14.53 we have then that f is $(r + 1)$-times differentiable.

Using mathematical induction we have that $S = \mathbb{N}$ proving the theorem. □

Theorem 14.66. *Let $\langle X, \|\|_X \rangle$ be a normed space, $\{\langle Y_i, \|\|_i \rangle\}_{i \in \{1, \ldots, n\}}$ a family of normed spaces and $\|\|$ the maximum norm defined on $\prod_{i \in \{1, \ldots, n\}} Y_i$ then given $U \subseteq X$, $V \subseteq \prod_{i \in \{1, \ldots, n\}} Y_i$ open sets and $r \in \mathbb{N}_0$ we have that $f: U \rightarrow \prod_{i \in \{1, \ldots, n\}} Y_i$ is C^r iff $\pi_i \circ f$ is C^r*

Proof. We have the following cases to consider:

1. **($r=0$)** This follows from 12.145.

2. **($r>0$)**

   a. **($\Rightarrow$)** If f is C^r then f is r-times differentiable, using the previous theorem we have then $\forall i\in\{1,\ldots,n\}$ that $\pi\circ f$ is r-times differentiable with $D^r(\pi_i\circ f)=\pi_i\circ D^r f$, as π_i and $D^r f$ is continuous we have that $\pi_i\circ f$ is C^r

   b. **($\Leftarrow$)** If $\forall i\in\{1,\ldots,n\}$ we have $\pi_i\circ f$ is C^r then $\pi_i\circ f$ is r-times differentiable so by the previous theorem we have that f is r-times differentiable and also by the previous theorem we must then have that $\pi_i\circ D^r f=D^r(\pi_i\circ f)$ proving that $\pi_i\circ D^r f$ is continuous and by 12.145 that $D^r f$ is continuous. So we have that f is C^r □

Lemma 14.67. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle,\langle Z,\|\|_Z\rangle$ be normed spaces, $U\subseteq X$ open, $V,W\subseteq Y$ open, $S\subseteq Z$ a open set, $f:U\to V$ a 1-times differentiable function [or C^1] and $g:W\to S$ a 1-times differentiable function [or C^1] [so using 2.32 $g\circ f$: $U\bigcap f^{-1}(W)\to S$ is a function] then $g\circ f:U\bigcap f^{-1}(W)\to S$ is 1-times differentiable on U [or C^1] and $D^1(g\circ f)=(\circ)\circ h$ where $(\circ):L(Y,Z)\times L(X,Y)\to L(X,Z)$ is defined by $(g,f)\to(\circ)(g,f)=g\circ f$ (see 12.196) is C^∞ and $h:U\to L(Y,Z)\times L(X,Y)$ is defined by $x\to h(x)=(D^1g(f(x)),D^1f(x))$ [is continuous if f,g is C^1]*

Proof. As f is 1-times differentiable on U and g is 1-times differentiable on V we have $\forall x\in U$ we that Df(x) exists and Dg($f(x)$) exists, using the chain rule (see 14.27) we have that $g\circ f$ is differentiable at x with $D^1(g\circ f)(x)=D(g\circ f)(x)=\mathrm{Dg}(f(x))\circ\mathrm{Df}(x)=D^1g(f(x))\circ D^1f(x)$. Using 12.196 we have that $(\circ)$ is continuous and bi-linear so it is by 14.63 C^∞. As $\pi_1\circ h=(D^1g)\circ f$ and and $\pi_2\circ h=D^1f$ we have by 12.145 that h is continuous, if g is C^1 (so that D^1g is continuous) and f is C^1 (so that D^1f is continuous). Finally $\forall x\in U$ we have $((\circ)\circ h)(x)=(\circ)(h(x))=(\circ)(D^1g(x),D^1f(x))=D^1g(f(x))\circ D^1f(x)\Rightarrow D^1(g\circ f)=(\circ)\circ h$ which is continuous if f,g is C^1 [so that h is continuous]. □

Theorem 14.68. (General Chain Rule) *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle,\langle Z,\|\|_Z\rangle$ be normed spaces, $U\subseteq X$ open, $V\subseteq Y$ open, $n\in\mathbb{N}$, $f:U\to Y$ a n-times differentiable function [or C^n] with $f(U)\subseteq V$ and $g:V\to Z$ a n-times differentiable function [or C^n] then $g\circ f:U\to Z$ is n-times differentiable [or C^n].*

Remark 14.69. If $n=0$ then from the fact that f,g is C^0 we have that f,g are continuous and thus that $g\circ f$ is continuous and thus $g\circ f$ is C^0.

Proof. We prove this by induction on n so let $S=\{n\in\mathbb{N}|$ if f: $U\to Y$ is $n-$ times differentiable [or C^n], $f(U)\subseteq V$, g: $V\to Z$ is $n-$ times differentiable[or C^n] then $g\circ f$ is times differentiable [or C^n]$\}$ then:

1. If $n=1$ then by the previous lemma we have $1\in S$.

2. If $n \in S$ assume then that f is $(n+1)$-times differentiable on U and g is $(n+1)$-times differentiable on V. Now $\pi_1 \circ h = D^1 g \circ f$ is the composition of two n-times differentiable functions [or C^n] so that by the fact that $n \in S$ we have that $\pi_1 \circ h$ is n-times differentiable [or C^n], likewise we have that $\pi_2 \circ h = D^1 f$ is n-times differentiable [or C^n] so that by 14.65 [or 14.66] we have that h is n-times differentiable [or C^n]. We have then by the previous theorem that $D^1(g \circ f) = (\circ) \circ h \underset{n \in S}{\Rightarrow} D^1(g \circ f)$ is n-times differentiable [or C^n] so that by 14.53 [or 14.54] we have that $g \circ f$ is $(n+1)$-times differentiable [or C^{n+1}] proving that $n+1 \in S$

Mathematical induction proves the theorem. □

Definition 14.70. *Let $n \in \mathbb{N}$ $\{X_i\}_{i \in \{1,\ldots,n+1\}}$ be a finite family of sets and $\{f_i\}_{i \in \{1,\ldots,n\}}$ functions from $X_i \rightarrow X_{i+1}$ then we define $f_n \circ \ldots \circ f_1: X_1 \rightarrow X_{n+1}$ recursively by*

1. *If $n = 1$ then $f_1 \circ \ldots \circ f_1 = f_1$*
2. *If $n > 1$ then $f_n \circ \ldots \circ f_1 = f_n \circ (f_{n-1} \circ \ldots \circ f)$*

By mathematical induction we can then generalize the above theorem

Theorem 14.71. *Let $n, k \in \mathbb{N}$, $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1,\ldots,n+1\}}$ be normed spaces, $\{U_i\}_{i \in \{1,\ldots,n\}}$ sets with $\forall i \in \{1,\ldots,n\}$ $U_i \subseteq X_i \wedge U_i$ open, $\{f_i: U_i \rightarrow X_{i+1}\}_{i \in \{1,\ldots,n\}}$ a family of functions with $\forall i \in \{1,\ldots,n-1\}$* that *$f_i(U_i) \subseteq U_{i+1}$ and $\forall i \in \{1,\ldots,n\}$ that f_i is r-times differentiable [or C^r] then $f_k \circ \ldots \circ f_1$ is n-times differentiable [or C^n]*

Proof. Let $S = \{k \in \mathbb{N} | \{\langle X_i, \|\|_i\rangle\}_{i \in \{1,\ldots,n+1\}}$ normed spaces $\{U_i\}_{i \in \{1,\ldots,n\}}$ sets with $\forall i \in \{1,\ldots,n\}$ $U_i \subseteq X_i$ open, $\{f_i: U_i \rightarrow X_{i+1}\}_{i \in \{1,\ldots,n\}}$ a family of functions with $\forall i \in \{1,\ldots,n\}$ $f_i(U_i) \subseteq U_{i+1}$ then $f_k \circ \ldots \circ f_1$ is $n-$times differentiable [or C^n]$\}$ then we have

1. If $k = 1$ then $f_1 \circ \ldots \circ f_1 = f_1$ so that $1 \in S$
2. If $k \in S$ then $f_{k+1} \circ \ldots \circ f_1 = f_{k+1} \circ (f_k \circ \ldots \circ f_1)$ which is n-times differentiable [or C^n] as f_k is n-times differentiable [or C^k] and $(f_k \circ \ldots \circ f_1)$ is n-times differentiable [or C^n] [as $k \in S$]. So we have $k+1 \in S$ □

Lemma 14.72. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces then the evaluation function* $\mathrm{ev}: L(X,Y) \times X \rightarrow Y$ *defined by $(L, v) \rightarrow \mathrm{ev}(L, v) = L(v)$ is bi-linear and continuous hence of class C^∞ (see 14.63) and also by 14.63* $D^1\mathrm{ev}(L, v) = \mathrm{ev}(L, *) + \mathrm{ev}(*, v)$

Proof.

1. **(bi-linearity)**
   a. If $\alpha, \beta \in \mathbb{K}$, $v \in X$ and $L_1, L_2 \in L(X,Y)$ then $\mathrm{ev}(\alpha \cdot L_1 + \beta \cdot L_2, v) = (\alpha \cdot L_1 + \beta \cdot L_2)(v) = \alpha \cdot L_1(v) + \beta \cdot L_2(v) = \alpha \cdot \mathrm{ev}(L_1, v) + \beta \cdot \mathrm{ev}(L_2)$
   b. If $\alpha, \beta \in \mathbb{K}, v_1 v_2 \in X$ and $L \in L(X,Y)$ then $\mathrm{ev}(L, \alpha \cdot v_1 + \beta \cdot v_2) = L(\alpha \cdot v_1 + \beta \cdot v_2) = \alpha \cdot L(v_1) + \beta \cdot L(v_2) = \alpha \cdot \mathrm{ev}(L, v_1) + \beta \cdot \mathrm{ev}(L, v_2)$

2. **(continuity)** $\|\mathrm{ev}(L,v)\|_Y \leqslant \|L\| \cdot \|v\|_X \leqslant 1 \cdot (\|L\| \cdot \|v\|)$ proving that ev is multilinear with $\|\mathrm{ev}\| \leqslant 1$ □

Lemma 14.73. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces, $U \subseteq X$ open, $n \in \mathbb{N}, n > 1$ and $f: U \to Y$ a n-times differentiable function [or f is C^n] then given $h \in X$ we have that $D^1 f(*)(h): U \to Y$ defined by $x \to D^1 f(x)(h)$ is $(n-1)$-times differentiable (or is C^{n-1}], further $\forall k \in X$ we have $D^1(D^1 f(*)(h))(x)(k) = D^2 f(x)(h)(k)$.*

Proof.

1. Given $v \in X$ define $\varphi_v: U \to L(X,Y) \times X$ by $x \to (D^1 f(x), v)$ then we have that $\pi_1 \circ \varphi_v = D^1 f$ which is $(n-1)$-times differentiable, $\pi_2 \circ \varphi_v = C_{U,v}$ which is C^∞ (see 14.60) so that by 14.65 φ_v is $(n-1)$-times differentiable [or by 14.66 φ_v is C^{n-1}]. Further by 14.65 we have that $\forall x \in U \vDash D^1 \varphi = (D(\pi_1 \circ \varphi_v), D(\pi_2 \circ \varphi_v)) = (D^2 f, C_{U,0})$

2. $\forall h \in X$, $\forall x \in U$ we have $D^1 f(*)(h)(x) = D^1 f(x)(h) = \mathrm{ev}(D^1 f(x), h) = \mathrm{ev}(\varphi_h(x)) = (\mathrm{ev} \circ \varphi_h)(x) \Rightarrow D^1 f(*)(h) = \mathrm{ev} \circ \varphi_h$ so that by the chain rule (see 14.68) we have that $D^1 f(*)(h)$ is $(n-1)$-times differentiable [or C^{n-1}]. Further $D^1(D^1 f(*)(h))(x) = D(D^1 f(*)(h))(x) = D((\mathrm{ev} \circ \varphi_h)(x)) \underset{14.27}{=} D\,\mathrm{ev}(\varphi_h(x)) \circ D\varphi_h(x) = (\mathrm{ev}(\pi_1(\varphi_h(x)), *) + \mathrm{ev}(*, \pi_2(\varphi_h(x)))) \circ (D^2 f(x), C_{U,0}) \Rightarrow D^1(D^1 f(*)(h))(x)(k) = (\mathrm{ev}(\pi_1(\varphi_h(x)), *) + \mathrm{ev}(*, \pi_2(\varphi_h(x))))(D^2 f(x), C_{U,0})(k) = (\mathrm{ev}(\pi_1(\varphi_h(x)), *) + \mathrm{ev}(*, \pi_2(\varphi_h(x))))(D^2 f(x)(k), 0) = \mathrm{ev}(D^1 f(x), 0) + \mathrm{ev}(D^2 f(x)(k), h) = D^1 f(x)(0) + D^2 f(x)(k)(h) = D^2 f(x)(k)(h)$ proving that $D^1(D^1 f(*)(h))(x)(k) = D^2 f(x)(k)(h)$ □

Lemma 14.74. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$be normed space, $n \in \mathbb{N}$, $f: U \to Y$ a n-times differentiable function [or C^n function] then $\Delta f: X \times X \to Y$ defined by $(x, h) \to \Delta f(x,h) = D^1 f(x)(h)$ is $(n-1)$-times differentiable [or C^{n-1}]*

Proof.

1. Define $\varphi: U \times X \to L(X,Y) \times X$ by $(x,v) \to (D^1 f(x), v)$ then as $\pi_1 \circ \varphi = D^1 f$ is by 14.53 [or 14.54] we have that $\pi_1 \circ \varphi$ is $(n-1)$-times differentiable [or C^{n-1}] and $\pi_2 \circ \varphi = 1_X$ which is of class C^∞ we have that φ is $(n-1)$-times differentiable [or C^{n-1}]

2. $(\mathrm{ev} \circ \varphi)(x,h) = \mathrm{ev}(\varphi(x,h)) = \mathrm{ev}(D^1 f(x), v) = D^1 f(x)(v) = \Delta f(x,v) \Rightarrow \mathrm{ev} \circ \varphi = \Delta f$ proving as φ is $(n-1)$-times differentiable [or C^{n-1}] and ev is C^∞ we have by the generalized chain rule that $\Delta f$is $(c-1)$ differentiable [or of class C^n] □

Theorem 14.75. *Let $\langle X_i, \|\|_i\rangle_{i \in \{1,\ldots,n\}}$, $\langle Y, \|\|_Y\rangle$be normed spaces, $X = \prod_{i \in \{1,\ldots,n\}} X_i$ be a the product space, $\langle X, \|\|\rangle$ a normed space with maximum norm (see 12.79) on X, $U \subseteq X$ open, $n \in \mathbb{N}$ [or $n \in \mathbb{N}_0$] and $f: U \to Y$ is n-times differentiable [or C^n] then $f^{(i)}: \pi_i(U) \to Y$ defined by $f \circ (x_1, \ldots, x_{i-1}, *, x_{i+1}, \ldots, x_n)_{|\pi_i(U)}$ is n-times differentiable [or C^n] with $D_i f = D^1 f^{(i)} \circ \pi_i$ and further $D_i f$ is $(n-1)$-times differentiable.*

Proof. First if $n=0$ then as $(x_1,\ldots,x_{i-1},*,x_{i+1},\ldots,x_n)$ is continuous and f is C^0 thus continuous we have that $f^{(i)}$ is C^0.

Second if $n>0$ then using 14.24 we have that $D(x_1,\ldots,x_{i-1},*,x_{i+1},\ldots,x_n)(t)=(0_i,\ldots,0_{i-1},*,0_{i+1},\ldots,0_n)$ is a linear continuous function, so C^∞ and thus by 14.53 we have that $(x_1,\ldots,x_{i-1},*,x_{i+1},\ldots,x_n)$ is of class C^∞. Using the generalized chain rule (see 14.68) we have then that $f^{(i)}$ is n-times differentiable [or C^n]. Also using 14.30 we have by definition that $D_i f(x)=Df^{(i)}(\pi_i(x))=D^1 f^{(i)}(\pi_i(x))\Rightarrow D_i f=D^1 f^{(i)}\circ\pi_i$. Which as π_i (being linear and continuous) is C^∞ means by 14.53 that $D_i f$ is $(n-1)$-times differentiable [or by 14.54 $D_i f$ is C^{n-1}]. □

Theorem 14.76. *Let $\langle X,\|\|_X\rangle$ be a normed space, $\langle\mathbb{K},||\rangle$ be $\mathbb{K}$ equipped with the canonical norm, $U\subseteq X$ and $n\in\mathbb{N}$ [or $n\in\mathbb{N}_0$] and $f,g:U\to\mathbb{K}$ are functions n-times differentiable [or of class C^n] then $f\cdot g:U\to\mathbb{K}$ is n-times differentiable [or of class C^r]*

Proof. First if $n=0$ then as f,g are C^0 and thus continuous we have by 12.172 that $f\cdot g$ is continuous and thus C^0. Second if $n>0$ defined then $\varphi:U\to\mathbb{K}\times\mathbb{K}$ by $x\to\varphi(x)=(f(x),g(x))$ then as $\pi_1\circ\varphi=f$ and $\pi_2\circ\varphi=g$ we have that φ is n-times differentiable [or C^n] (see 14.65 [or 14.66]). As $(\cdot):\mathbb{K}\times\mathbb{K}\to\mathbb{K}$ defined by $(x,y)\to x\cdot y$ is trivially bi-linear and continuous we have by 14.63 that $(\cdot)$ is C^∞. Now as $((\cdot)\circ\varphi)(x)=(\cdot)(\varphi(x))=(\cdot)(f(x),g(x))=f(x)\cdot g(x)$ we have $(\cdot)\circ\varphi=f\cdot g$ we have by the generalized chain rule that $f\cdot g$ is n-times differentiable [or C^n]. □

Theorem 14.77. *Let $n\in\mathbb{N}$, $\langle\mathbb{R}^n,\|\|\rangle$ be the set of real tupple's together with the maximum norm defined on $||$, $\langle Y,\|\|_Y\rangle$ be a Banach space, $U\subseteq\mathbb{R}^n$ open and $U\subseteq\mathbb{R}^n$ open and $f:U\times[a,b]\to Y$ (where $a<b$) and f a function satisfying the following conditions*

1. *f is a continuous function*
2. *$\forall t\in[a,b]$ the function $f(*,t):U\to Y$ defined by $x\to f(*,t)(x)=f(x,t)$ is differentiable of class $C^r, r\geqslant 1$*
3. *$\forall i\in\{1,\ldots,r\}$ the function $D^i f(*,*)(*):U\times[a,b]\to L^i(\mathbb{R}^n;Y)$ defined by $(x,t)\to(D^i f(*,*)(*))(x,t)=D^i f(*,t)(x)$ is continuous.*

then if we define $\forall x\in U$ $f(x,*):[a,b]\to Y$ by $t\to f(x,*)(t)=f(x,t)$ we have that the function $F(x):U\to Y$ defined by $x\to F(x)=\int_a^b f(x,*)$ is defined and of class C^r. Further $D^r F(x)=\int_a^b D^r f(*,*)(x)$ [where $D^r f(*,*)(x):[a,b]\to L^r(\mathbb{R}^n;Y)$ with $t\to(D^r f(*,*)(x))(t)=D^r f(*,t)(x)$]. A easier way to note this is by the classic notation of the Newton integral as $D^r F(x)=\int_a^b D^r f(*,t)(x)\cdot\mathrm{dt}$.

Proof. The proof is done in different stages

1. First we prove that $\forall x\in U$ $F(x)=\int_a^b f(x,*)$ is well defined. From the continuity of f and 12.146 we have that $\forall x\in U$ that $f(x,*)$ is continuous, So the integral $\int_a^b f(x,*)$ is well defined.

2. The rest of the proof is done by induction so let $S = \{r \in \mathbb{N} | \text{if } f \text{ is function satisfying (1), (2), (3) of the theorem then} \forall x \in U F(x) \text{ is defined, of class } C^r \text{ and } D^r F(x) = \int_a^b D^r f(*, *)(x)\}$ then we have to prove:

   a. **(Base case $r = 1$)** First using 12.352 we have that $L^1(\mathbb{R}^n; Y) = L(\mathbb{R}^n; Y)$ is a Banach space. Further $\forall x \in U$ we have that $f_1^{(1)}: [a, b] \to L^1(\mathbb{R}; Y)$ defined by $t \to f_1^{(1)}(t) = D^1 f(*, t)(x)$ is continuous by (3) and12.146. So the integral $\int_a^b f_1^{(1)} = \int_a^b D^1 f(*, *)(x) \in L^1(\mathbb{R}^n; Y)$ exists and is well defined. Now given $h \in U_x$ we have if we define

$$\Delta(x, h) = F(x + h) - F(x) - \left(\int_a^b D^1 f(*, *)(x)\right)(h) \tag{14.25}$$

   that

$$\begin{aligned}\Delta(x, h) &\underset{12.436}{=} F(x + h) - F(x) - \int_a^b [D^1 f(*, *)(x)](\cdot)(h) \\ &= \int_a^b f(x + h, *) - \int_a^b f(x, *) - \int_a^b [D^1 f(*, *)(x)](\cdot)(h) \\ &= \int_a^b [f(x + h, *) - f(x, *) - [D^1 f(*, *)(x)](\cdot)(h)]\end{aligned}$$

   giving

$$\Delta(x, h) = \int_a^b [f(x + h, *) - f(x, *) - [D^1 f(*, *)(x)](\cdot)(h)] \tag{14.26}$$

   As given $t \in [a, b]$ we have by (2) that $f(*, t)$ is C^1 thus differentiable at every $x \in U$ there exists by 14.8 a ε-function $\varepsilon_{x,t}: U_x \to Y$ such that

$$\begin{aligned}\varepsilon_{x,t}(h) \cdot \|h\| &= f(*, t)(x + h) - f(*, t)(x) - \mathrm{Df}(*, t)(x)(h) \\ &= f(x + h, t) - f(x, t) - D^1 f(*, t)(x)(h) \\ &= f(x + h, t) - f(x, t) - (D^1 f(*, *)(x))(t)(h) \\ &= f(x + h, *)(t) - f(x, *)(t) - [(D^1 f(*, *)(x))(\cdot)(h)](t) \\ &= [f(x + h, *) - f(x, *) - (D^1 f(*, *)(x))(\cdot)(h)](t)\end{aligned}$$

   and

$$\varepsilon_{x,t}(0) = 0$$

   This defines using the Axiom of Choice a function $\varepsilon_x: [a, b] \to \mathcal{C}(U_x, Y)$ such that $\forall t \in [a, b]$ we have $\varepsilon_x(t)(h) \cdot \|h\| = [f(x + h, *) - f(x, *) - [D^1 f(*, *)(x)](\cdot)(h)](t)$ or $(\|h\| \cdot \varepsilon_x(\cdot)(h))(t) = \varepsilon_x(t)(h) \cdot \|h\| = [f(x + h, *) - f(x, *) - [D^1 f(*, *)(x)](\cdot)(h)](t)$ giving

$$\|h\| \cdot \varepsilon_x(\cdot)(h) = f(x + h, *) - f(x, *) - [D^1 f(*, *)(x)](\cdot)(h)\,, \varepsilon_x(.)(0) = 0 \tag{14.27}$$

so that using 14.26 we have

$$\Delta(x,h) = \int_a^b \|h\| \cdot \varepsilon_x(\cdot)(h) \tag{14.28}$$

Next define $\varphi_x: U_x \times [a,b] \to Y$ defined by $(h,t) \to \varphi_x(h,t) = D^1 f(*,t)(x)(h) = \mathrm{Df}(*,t)(x)(h)$ and prove that φ_x is continuous. So if (h',t'), $(h,t) \in U_x \times [a,b]$ then

$$\begin{aligned}
\|\varphi_x(h',t') - \varphi_x(h,t)\|_Y &= \|\mathrm{Df}(*,t')(x)(h') - \mathrm{Df}(*,t)(x)(h)\|_Y \\
&= \|\mathrm{Df}(*,t')(x)(h') - \mathrm{Df}(*,t)(x)(h') + \mathrm{Df}(*, \\
&\quad t)(x)(h') - \mathrm{Df}(*,t)(x)(h)\|_Y \\
&\leqslant \|\mathrm{Df}(*,t')(x)(h') - \mathrm{Df}(*,t)(x)(h')\|_Y + \\
&\quad \|\mathrm{Df}(*,t)(x)(h') - \mathrm{Df}(*,t)(x)(h)\|_Y \\
&\leqslant \|\mathrm{Df}(*,t)(x) - \mathrm{Df}(*,t)(x)\|_{L^1(\mathbb{R}^n;Y)} \cdot \|h'\| + \\
&\quad \|\mathrm{Df}(*,t)(x)(h') - \mathrm{Df}(*,t)(x)(h)\|_Y \\
&\leqslant \|\mathrm{Df}(*,t)(x) - \mathrm{Df}(*,t)(x)\|_{L^1(\mathbb{R}^n;Y)} \cdot \|h'\| + \\
&\quad \|\mathrm{Df}(*,t)(x)(h'-h)\|_Y \\
&\leqslant \|\mathrm{Df}(*,t)(x) - \mathrm{Df}(*,t)(x)\|_{L^1(\mathbb{R}^n;Y)} \cdot \|h'\| + \\
&\quad \|\mathrm{Df}(*,t)(x)\|_{L^1(\mathbb{R}^n;Y)} \cdot \|h-h'\|
\end{aligned}$$

so we have

$$\|\varphi_x(h',t') - \varphi_x(h,t)\|_Y \leqslant \|\mathrm{Df}(*,t)(x) - \mathrm{Df}(*,t)(x)\|_{L^1(\mathbb{R}^n;Y)} \cdot \|h'\| + \|\mathrm{Df}(*,t)(x)\|_{L^1(\mathbb{R}^n;Y)} \cdot \|h-h'\| \tag{14.29}$$

Now given $(h,t) \in U_x \times [a,b]$ and $\varepsilon > 0$. Then from the continuity of $D^1 f(*,*)(*)$ we have by 12.146 that $D^1 f(*,*)(x): [a,b] \to (D^1 f(*,*)(x))(t) = (D^1 f(*,*)(*))(x,t) = D^1 f(*,t)(x)$ is continuous. So that there exists a δ_1 such that if $|t-t'| < \delta_1$ then $\|D^1 f(*,t)(x) - D^1 f(*,t')(x)\|_{L^1(\mathbb{R}^n;Y)} = \|(D^1 f(*,*)(x))(t) - (D^1 f(*,*)(x)(t'))\|_{L^1(\mathbb{R}^n;Y)} < \frac{\varepsilon}{2 \cdot (\|h\|+1)}$ so that

$$|t-t'| < \delta_1 \Rightarrow \|D^1 f(*,t)(x) - D^1 f(*,t')(x)\|_{L^1(\mathbb{R}^n;Y)} < \frac{\varepsilon}{2 \cdot (\|h\|+1)} \tag{14.30}$$

Take now $\delta = \min\left(1, \frac{\varepsilon}{2 \cdot (\|\mathrm{Df}(*,t)(x)\|_{L^1(\mathbb{R}^n;Y)} + 1)}, \delta_1\right)$ then if we have $\|h - h'\| < \delta \Rightarrow \|D^1 f(*,t)(x)\|_{L^1(\mathrm{Rr}^n;Y)} \cdot \|h-h'\| \leqslant \|D^1 f(*,t)(x)\|_{L^1(\mathbb{R}^n;Y)} \cdot \frac{\varepsilon}{2 \cdot (\|\mathrm{Df}(*,t)(x)\|_{L^1(\mathbb{R}^n;Y)} + 1)} < \frac{\varepsilon}{2}$ so that

$$\|h-h'\| < \delta \Rightarrow \|D^1 f(*,t)(x)\|_{L^1(\mathrm{Rr}^n;Y)} \cdot \|h-h'\| < \frac{\varepsilon}{2} \tag{14.31}$$

As $\|h'\| \leqslant \|h' - h + h\| \leqslant \|h\| + \|h' - h\| \leqslant \|h\| + 1$ we have

$$\|h - h'\| < \delta \Rightarrow \|h'\| \leqslant \|h\| + 1 \tag{14.32}$$

So if $\max(\|h - h'\|, |t - t'|) = \|(h - h', t - t')\| = \|(h, t) - (h', t')\| < \delta$ then using 14.29,14.30 we have that

$$\begin{aligned}
\|\varphi_x(h,t) - \varphi_x(h',t')\|_Y &< \frac{\varepsilon}{2 \cdot (\|h\| + 1)} \cdot \|h'\| + \frac{\varepsilon}{2} \\
&\leqslant \frac{\varepsilon}{2 \cdot (\|h\| + 1)} \cdot (\|h\| + 1) + \frac{\varepsilon}{2} \\
&\leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon
\end{aligned}$$

proving that φ_x is indeed continuous. Define now $\varepsilon'_x: U_x \times [a, b] \to Y$ where $\varepsilon'_x(h, t) = \|h\| \cdot \varepsilon_x(t)(h) = (\|h\| \cdot \varepsilon_x(\cdot)(h))(t) \underset{14.27}{=} [f(x + h, *) - f(x, *) - [D^1 f(*, *)(x)](\cdot)(h)](t) = f(x + h, t) + f(x, t) + D^1 f(*, t)(x)(h) = f(x + h, t) + f(x, t) + \varphi(x, t)$. As φ is continuous, f is continuous and $(h, t) \to (x + h, t)$ is continuous we have that ε'_x is a continuous function. So we have proved that

$$\varepsilon'_x: U_x \times [a, b] \to Y \text{ defined by } \varepsilon'_x(h, t) = (\|h\| \cdot \varepsilon_x(\cdot)(h)(t)) \text{ is continuous} \tag{14.33}$$

Define now $\xi_x: U_x \times [a, b] \to Y$ defined by $(h, t) \to \xi_x(t)(h) = (\varepsilon_x(\cdot)(h))(t) = \varepsilon_x(t)(h)$ then we prove that ξ_x is continuous at $U_x \times [a, b]$. So let $(h, t) \in U_x \times [a, b]$ and $\varepsilon > 0$ then we have two cases to consider:

i. $\boldsymbol{((h, t) \in (U_x \setminus \{0\}) \times [a, b])}$. As $h \neq 0$ we have $\|h\| > 0$ so that there exists a $\lambda > 0$ such that $\|h\| > \lambda > 0$. Take $\delta_1 = \frac{\|h\| - \lambda}{2}$ then if $h' \in U'_x$ with $\|h - h'\| < \delta_1$ then $0 < \lambda < \|h'\|$ [for if $\|h'\| \leqslant \lambda$ then $\|h\| \leqslant \|h - h'\| + \|h'\| < \frac{\|h\| - \lambda}{2} + \|h'\| \leqslant \frac{\|h\| - \lambda}{2} + \lambda = \frac{\|h\| - \lambda + 2 * \lambda}{2} = \frac{\|h\|}{2} + \frac{\lambda}{2} \Rightarrow \|h\| - \frac{\|h\|}{2} < \frac{\lambda}{2} \Rightarrow \frac{\|h\|}{2} < \frac{\lambda}{2} \Rightarrow \|h\| < \lambda < ||h| \Rightarrow \|h\| < \|h\|$ a contradiction]. So given (h', t') with $\|h - h'\| < \delta_1$ and thus $\|h'\| > 0$ then we have

$$\begin{aligned}
\|\xi_x(h,t) - \xi_x(h',t')\|_Y &= \|\varepsilon_x(t)(h) - \varepsilon_x(t')(h')\|_Y \\
&\underset{\|h\|, \|h'\| \geqslant 0}{=} \left\| \frac{\|h\| \cdot \varepsilon_x(t)(h)}{\|h\|} - \frac{\|h'\| \cdot \varepsilon_x(t')(h')}{\|h'\|} \right\|_Y \\
&= \left\| \frac{\varepsilon'_x(h,t)}{\|h\|} - \frac{\varepsilon'(h',t')}{\|h'\|} \right\|_Y \\
&= \frac{1}{\|h\| \cdot \|h'\|} \cdot \| \|h'\| \cdot \varepsilon'_x(h,t) - \|h\| \cdot \varepsilon'_x(h',t') \|_Y
\end{aligned}$$

$$\begin{aligned} &\leqslant \frac{1}{\|h\| \cdot \|h'\|} \| \|h'\| \cdot \varepsilon'_x(h,t) - \|h\| \cdot \varepsilon'_x(h,t)\|_Y + \| \|h\| \cdot \varepsilon'_x(h,t) - \|h\| \cdot \varepsilon'_x(h',t')\|_Y \\ &= \frac{1}{\|h\| \cdot \|h'\|} \cdot [|\|h'\| - \|h\|| \cdot \|\varepsilon'_x(h,t)\|_Y + \|h\| \cdot \|\varepsilon'_x(h,t) - \varepsilon'_x(h',t')\|_Y] \end{aligned}$$

proving that

$$\|\xi_x(h,t) - \xi_x(h',t')\|_Y \leqslant \frac{1}{\|h\|h'} \cdot [(\|h'\| - \|h\|) \cdot \|\varepsilon'_x(h,t)\|_Y + \|h\| \cdot \|\varepsilon'_x(h,t) - \varepsilon'_x(h',t')\|_Y] \tag{14.34}$$

Now by continuity of ε'_x we can choose a $\delta_2 > 0$ such that if $\|(h,t) - (h',t')\| < \delta_2$ then $\|\varepsilon'_x(h,t) - \varepsilon'_x(h',t')\|_Y < \frac{\varepsilon \cdot \lambda}{2} \Rightarrow \frac{1}{\|h\| \cdot \|h'\|} \cdot \|h\| \cdot \|\varepsilon'_x(h,t) - \varepsilon'_x(h,t)\|_Y \leqslant \frac{\varepsilon \cdot \lambda}{2 \cdot \|h'\|} < \frac{\varepsilon}{2}$ [as $\lambda < \|h'\|$] so that by 14.34 we have if $\max(\|h - h'\|, |t - t'|) = \|(h,t) - (h',t')\| < \min(\delta_1, \delta_2)$ then

$$\|\xi_x(h,t) - \xi_x(h',t')\|_Y < \frac{1}{\|h\| \cdot \|h'\|} \cdot |\|h'\| - \|h\|| \cdot \|\varepsilon'_x(h,t)\|_Y + \frac{\varepsilon}{2} \tag{14.35}$$

Take now $\delta = \min\left(\delta_1, \delta_2, \frac{\varepsilon \cdot \lambda^2}{2 \cdot (\|\varepsilon'_x(h,t)\|_Y + 1)}\right)$ then if $\max(\|h - h'\|, |t - t'|) = \|(h,t) - (h',t')\| < \delta$ we have by 14.35 that

$$\begin{aligned} \|\xi_x(h,t) - \xi_x(h',t')\|_Y &< \frac{1}{\|h\| \cdot \|h'\|} \cdot |\|h'\| - \|h\|| \cdot \|\varepsilon'_x(h,t)\|_Y + \frac{\varepsilon}{2} \\ &\leqslant \frac{1}{\|h\| \cdot \|h'\|} \cdot \|h - h'\| \cdot \|\varepsilon'_x(h,t)\|_Y + \frac{\varepsilon}{2} \\ &< \frac{\varepsilon \cdot \lambda^2 \cdot \|\varepsilon'_x(h,t)\|_Y}{\|h\| \cdot \|h'\| \cdot 2 \cdot (\|\varepsilon'_x(h,t)\|_Y + 1)} + \frac{\varepsilon}{2} \\ &< \frac{\varepsilon \cdot \lambda^2}{2 \cdot \|h\| \cdot \|h'\|} + \frac{\varepsilon}{2} \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

proving finally that ξ_x is continue at (h,t).

ii. $\boldsymbol{((h,t) = (0,t) \in \{0\} \times [a,b])}$ then by 14.27 we have $\forall t \in [a,b] \vDash \xi_x(0,t) = \varepsilon_x(t)(h) = 0$, also if $t \in [a,b]$ we have by continuity of $\varepsilon_x(t)$ at 0 that $\exists \delta_t > 0$ such that if $\|h' - 0\| = \|h'\| < \delta_t$ then $\|\varepsilon_x(t)(h)\|_Y = \|\varepsilon_x(t)(0) - \varepsilon_x(t)(h)\|_Y < \varepsilon$ so that if $(h',t') \in U_x \times [a,b]$ with $\max(\|h'\|, |t - t'|) = \|(0 - h, t - t')\| < \delta$ then $\|\xi_x(0,t) - \xi_x(h',t')\|_Y = \|\varepsilon_x(t)(0) - \varepsilon_x(t)(h)\|_Y < \varepsilon$ proving continuity at $(0,t)$.

Given now a $x\in U$ then as 14.3 we can find a $\delta>0$ such that $0\in B_{\|\|}(0,2\cdot\delta)\subseteq U_x$ so that $0\in B_{\|\|}(0,\delta)\subseteq\overline{B_{\|\|}}(0,\delta)\subseteq B_{\|\|}(0,2\cdot\delta)\subseteq U_x$. The set $K=\overline{B_{\|\|}}(x,\delta)\times[a,b]$ is compact because of 12.288 (corollary of Tychonoff's), 12.257, 12.290 (Heine Borel) and 12.56 ($\overline{B_{\|\|}}(0,\delta)$ is closed). So by 12.258 we have that ξ_x is uniform continuous on K. So given a $\varepsilon>0$ there exists a $\delta'>0$ such that $\forall(h,t)\in K$ we have if $(h',t')\in K$ is such that $\|(h,t)-(h',t')\|<\delta'$ then $\|\xi_x(h,t)-\xi_x(h',t')\|_Y<\frac{\varepsilon}{b-a}$. More specifically $\forall t\in[a,b]$ we have if $\|h-0\|=\|h\|<\min(\delta.\delta')$ then $\|(h,t)-(0,t)\|=\|(h,0)\|<\delta'$ then $\|\xi_x(h,t)-\xi_x(0,t)\|_Y<\frac{\varepsilon}{b-a}$ from which it follows that [if $\xi_x(h,*)\colon[a,b]\to Y$ is defined by $\xi_x(h,*)(t)=\xi_x(h,t)$]

$$
\begin{aligned}
\left\|\int_a^b\xi_x(h,*)\right\|_Y &\leqslant \int_a^b\|\xi_x(h,*)\|_Y \text{ (see 12.433)}\\
&< (b-a)\cdot\frac{\varepsilon}{b-a} \text{ (see 12.433 again)}\\
&= \varepsilon
\end{aligned}
$$

this together with

$$
\begin{aligned}
\int_a^b\xi_x(0,t) &= \int_a^b C_0\\
&\underset{12.433}{=} (b-a)\cdot 0=0
\end{aligned}
$$

proves that $\zeta\colon U_x\to Y$ defined by $h\to\zeta(h)=\int_a^b\xi_x(h,*)$ is a ε-function. As $\xi(h,*)(t)=\varepsilon_x(t)(h)=\varepsilon_x(\cdot)(h)$ we have by 14.25, 14.28 and 12.432 that $F(x+h)-F(x)-\left(\int_a^b D^1f(*,*)(x)\right)(h)=\|h\|\cdot\int_a^b\xi_x(h,\cdot)$ a ε-function, proving that F is differentiable for every $x\in U$ with $\mathrm{DF}(x)=\int_a^b D^1f(*,*)(x)$ so that F is differentiable on U with $D^1F(x)=\int_a^b D^1f(*,*)(x)$. Now to prove that F is C^1 we must prove that $D^1F\colon U\to L(\mathbb{R},Y)$ is continuous. Let $x\in U$ then we can find a $\delta>0$ such that $x\in B_{\|\|}(x,\delta)\subseteq\overline{B_{\|\|}}(x,\delta)\subseteq B_{\|\|}(x,2\delta)\subseteq U$ then the set $K=\overline{B_{\|\|}}(x,\delta)\times[a,b]$ is compact because of 12.288 (corollary of Tychonoff's), 12.257, 12.290 (Heine Borel) and 12.56 ($\overline{B_{\|\|}}(0,\delta)$ is closed). From the continuity of $D^1f(*,*)(*)$ we have then that $D^1f(*,*)(*)$ is uniform continuous on K. So $\forall\varepsilon>0$ there exists a $\delta'>0$ such that if $(x,t),(x',t')\in K\times[a,b]$ with $\|(x,t)-(x',t')\|<\delta'$ then $\|D^1f(*,t)(x)-D^1f(*,t')(x')\|_{L^1(\mathbb{R}^n;Y)}=\|(D^1f(*,*)(*))(h,t)-(D^1f(*,*)(*))(h',t')\|_{L^1(\mathbb{R}^n;Y)}<\frac{\varepsilon}{b-a}$ or

$$
\text{If }\|(x,t)-(x',t')\|<\delta'\wedge(x,t),(x',t')\in K\times[a,b]\Rightarrow\|D^1f(*,t)(x)-D^1f(*,t')(x')\|_{L^1(\mathbb{R}^n;Y)}<\frac{\varepsilon}{b-a} \tag{14.36}
$$

So if we take $x' \in U$ such that $\|x-x'\| \leqslant \delta$ then $x, x' \in K, t \in [a,b]$ so that

$$\begin{aligned}
\|D^1F(x)-D^1F(x')\|_{L^1(\mathbb{R}^n;Y)} &= \left\| \int_a^b D^1f(*,*)(x) - \int_a^b D^1f(*,*)(x') \right\|_{L^1(R^n;Y)} \\
&\underset{12.432}{=} \left\| \int_a^b (D^1f(*,*)(x) - D^1f(*,*)(x')) \right\| \\
&\leqslant \int_a^b \|D^1f(*,*)(x) - D^1f(*,*)(x')\| \\
&< (b-a) \cdot \frac{\varepsilon}{b-a} \text{ (see 14.36 and 12.433)} \\
&= \varepsilon
\end{aligned}$$

proving that $D'F$ is continuous and thus finally that $1 \in S$, which is the end of the first induction step.

b. **(Induction step, let $r \in S$)** and assume that $\forall t \in [a,b] \vDash f(*,t)$ is C^{r+1} and $\forall i \in \{1,\ldots,r+1\}$ we have that $D^i f(*,*)(*)$ is continuous. Define then $g: U \times [a,b] \to L^r(\mathbb{R}^n;Y)$ by $(x,t) \to g(x,t) = D^r f(*,t)(x)$ so that $g = D^r f(*,*)(*)$ which by assumption is continuous. Then we have for g that $\forall t \in [a,b]$ that $g(*,t): U \to L^r(\mathbb{R}^n;Y)$ is defined by $x \to g(*,t)(x) = g(x,t) = D^r f(*,t)(x)$ which as $f(*,t)$ is C^{r+1} means that

$\forall t \in [a,b]$ we have that $g(*,t)$ is C^1 on U and $D^1 g(*,t)(x)$ exists $\forall x \in U$ (14.37)

where

$$g(*,t) = D^r f(*,t) \tag{14.38}$$

Also $D^1 g(*,*)(*): U \times [a,b] \to L(\mathbb{R}^n, L^r(\mathbb{R}^n;Y)) = L^{r+1}(\mathbb{R}^n;Y)$ defined by $(x,t) \to (D^1 g(*,*)(*))(x,t) = D^1 g(*,t)(x) = D^1(D^r f(*,t))(x) = D^{r+1} f(*,t)(x) = (D^{r+1} f(*,*)(*))(x,t)$ proving that

$D^1 g(*,*)(*) = D^{r+1} f(*,*)(*)$ which is continuus by the hypothese (14.39)

So g satisfies (1), (2) and (3) of the theorem for the case $r = 1$ and we can use then $1 \in S$ to have that

$$G: U \to L^r(\mathbb{R}^n;Y) \text{ defined by } x \to G(x) = \int_a^b g(x,*) \text{ is of } C^1 \tag{14.40}$$

and that

$$D^1 G(x) = \int_a^b D^1 g(*, *)(x). \tag{14.41}$$

Now $g(x, *): [a, b] \rightarrow L^r(\mathbb{R}^n; Y)$ is defined by $g(x, *)(t) = g(x, t) = D^r f(*, t)(x) = D^r f(*, *)(x)(t)$ so that $g(x, *) = D^r f(*, *)(x)$ so that

$$\begin{aligned} G(x) &\underset{14.40}{=} \int_a^b D^r f(*, *)(x) \\ &\underset{r \in S}{=} D^r F(x) \end{aligned}$$

or

$$G(x) = D^r F(x) \tag{14.42}$$

which as G is C^1 means that

$$F \text{ is } C^{n+1} \tag{14.43}$$

Now $(D^1 g(*, *)(x))(t) = D^1 g(*, t)(x) = (D^1 g(*, *)(*))(x, t) \underset{14.39}{=} (D^{r+1} f(*, *)(*))(x, t) = D^{r+1} f(*, t)(x) = (D^{r+1} f(*, *)(x))(t)$ proving that

$$D^1 g(*, *)(x) = D^{r+1} f(*, *)(x) \tag{14.44}$$

$$\begin{aligned} D^{r+1} F(x) &= D^1(D^r F)(x) \\ &\underset{14.42}{=} D^1 G(x) \\ &\underset{14.41}{=} \int_a^b D^1 g(*, *)(x) \\ &\underset{14.44}{=} \int_a^b D^{r+1} f(*, *)(x) \end{aligned}$$

proving

$$D^{r+1} F(x) = \int_a^b D^{r+1} f(*, *)(x) \tag{14.45}$$

Using 14.43and 14.45 we have then that $r + 1 \in S$.

by induction we have then that $S = \mathbb{N}$ which proves our little theorem. □

Theorem 14.78. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ and $\langle Z, \|\|_Z \rangle$ be normed spaces $U \subseteq X$ a open set $f: X \rightarrow Y$ a function that is C^n and $L: Y \rightarrow Z$ a linear continuous function then $L \circ f$ is C^n and $\forall x \in U$ we have $D^n(L \circ f)(x) = L \bullet_n D^n f(x)$ (see 12.214 for a definition of $\bullet_n$) (so we have $D^n(L \circ f)(x)(h_1: \ldots: h_n) = L(D^n f(x)(h_1: \ldots: h_n))$*

Proof. First as L is C^∞ and f is C^n on U we have by the Chain Rule (see 14.68) that $L \circ f$ is C^n. So the only thing left to prove is that $D^n(L \circ f)(x) = L \bullet_n D^n f(x)$ so take $S = \{r \in \mathbb{N} | \text{if } r \leqslant n \text{ and } x \in U \text{ then } D^r(L \circ f)(x) = L \bullet_r D^n f(x)\}$ then we have:

1. If $r = 1$ then as $D^1(L \circ f)(x) = D^1 L(f(x)) \circ D^1 f(x) = L \circ D^1 f(x) \underset{12.214}{=} L \bullet_1 D^1 f(x)$ proving that $r \in S$

2. Assume that $r \in S$ and that $r+1 \leqslant n$ then $D^r(L \circ f)$ is differentiable and $D^r(L \circ f)(x) = L \bullet_r D^r f(x) \underset{12.216}{=} \mathrm{ev}_L(D^r f(x)) = (\mathrm{ev}_L \circ D^r f)(x)$ so that $D^{r+1}(L \circ f)(x) = D(D^r(L \circ f))(x) = D(\mathrm{ev}_L \circ D^r f)(x) = \mathrm{Dev}_L(D^r f(x)) \circ D(D^r f)(x) \underset{\text{by } 12.216\, \mathrm{ev}_L \text{ is linear and continuous}}{=} \mathrm{ev}_L \circ D^{r+1} f(x) \underset{12.217}{=} L \bullet_{r+1} D^{r+1} f(x)$ proving that $r+1 \in S$. □

14.3 Differentiability on general sets

Definition 14.79. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be normed spaces $S \subseteq X$ and $f: S \to Y$ a function then f differentiable on S if there exists a U open such that $S \subseteq U \subseteq X$ and a $f^U: U \to Y$ that is differentiable on U. Also f is C^r if there exists a U open such that $S \subseteq U \subseteq X$ and a $f^U: U \to Y$ with $(f^U)_{|S} = f$ such that f^U is C^r. In other words f is C^r if it can be extended to a C^rfunction on a open set containing S.*

Theorem 14.80. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$, $\langle Z, \|\|_Z \rangle$ be normed spaces $S \subseteq X$, $T \subseteq Y$, $r \in \mathbb{N}_0$ and $f: S \to Y$, $g: T \to Z$ functions with $f(S) \subseteq T$ then $g \circ f$ is C^r on S.*

Proof. By definition there exists a U open in X and a V open in Y such that $S \subseteq U$, $T \subseteq U$ and C^r functions $f^U: U \to X$, $g^V: V \to Z$ such that $(f^U)_{|S} = f$ and $(g^V)_{|T} = g$. Take then $W = U \bigcap (f^U)^{-1}(V)$ which is open because f^U is continuous on V (being of class C^r). As $f^U(S) = f(S) \subseteq T \subseteq V \Rightarrow S \subseteq (f^U)^{-1}(V) \subseteq W \Rightarrow S \subseteq W$. Define now $f^W = (f^U)_{|W}: W \to Y$ then f^W is C^r (see 14.56) and as $S \subseteq W$ f^W is still a extension of f. Now $f^W(W) = f^U(W) \subseteq f^U((f^U)^{-1}(V)) \subseteq V$ so that by using the chain rule (14.68) we have that $f^W \circ g^V$ is C^r and as $(f^W \circ g^V)_{|S} = g \circ f$ we have that $g \circ f$ is C^r. □

Definition 14.81. *Let $\langle \mathbb{R}, \|\rangle$ be the real space with the canonical norm, $\langle Y, \|\|_Y \rangle$ a normed space, $[a, b] \subseteq \mathbb{R}, a < b$ and $f: [a, b] \to Y$ then we define the right,left derivative by*

1. *If $x \in [a, b[$ then $f'_+(x)$ is the right derivative of f at x if and only if $\forall \varepsilon > 0$ there exists a $\delta > 0$ such that for all h with $0 < h < \delta$ and $x + h \in [a, b]$ we have $\left\| \frac{f(x+h) - f(x)}{h} - f'_+(x) \right\|_Y < \varepsilon$*
2. *If $x \in]a, b]$ then $f'_-(x)$ is the left derivative of f at x if and only if $\forall \varepsilon > 0$ there exists a $\delta > 0$ such that for all h with $0 < h < \delta$ and $x - h \in [a, b]$ we have $\left\| \frac{f(x) - f(x-h)}{-h} - f'_-(x) \right\| < \varepsilon$*

Lemma 14.82. *Let $\langle \mathbb{R}, \|\rangle$ be the real space with the canonical norm, $[a, b] \subseteq \mathbb{R}$, $a < b$ and $f: [a, b] \to Y$ then*

1. *If $x \in [a, b[$ and $f'_+(x)$ if it exists is unique*
2. *If $x \in]a, b]$ and $f'_-(x)$ if it exists is unique*
3. *If $x \in]a, b[$ and $f'_-(x), f'_+(x)$ both exists and $f'_-(x) = f'_+(x)$ then $f'(x)$ exists and $f'(x) = f'_-(x) = f'_+(x)$. Also if $f'(x)$ exists then $f'_-(x), f'_+(x)$ exists and $f_+(x) = f_-(x)$.*

4. *If* $[a, b] \subseteq U$ *open and* $f^U: U \rightarrow Y$ *is a function with* $(f^U)_{|[a,b]} = f$ *for which* $(f^U)'(x)$ *exists* $\forall x \in U$ *then* $f'(x)$ *exists on* $]a, b[$ *with* $(f^U)'(x) = f'(x)$, $f'_+(a), f'_-(b)$ *exists with* $f'_+(a) = (f^U)'(a)$ *and* $(f^U)'(b) = f'_-(b)$

Proof.

1. Assume that there exists two different right derivatives d, d' at x define then $\varepsilon = \|d' - d\|_Y > 0$, then there exists $\delta, \delta' > 0$ such that if $0 < h < \delta$ and $x + h \in [a, b]$ that $\left\|\frac{f(x+h) - f(x)}{f} - d\right\|_Y < \frac{\varepsilon}{2}$ and if $0 < h < \delta'$ and $x - h \in [a, b]$ then $\left\|\frac{f(x+h) - f(x)}{h} - d'\right\|_Y < \frac{\varepsilon}{2}$. As $x \in [a, b[$ we have $a \leqslant x < b$ there exists a $\delta'' > 0$ such that $a \leqslant x + \delta'' < b$ take then $h = \min(\delta'', \delta', \delta)$ then if $0 < h < \delta, \delta'$ then $x + h \in [a, b]$ and $\|d' - d\|_Y < \left\|\frac{f(x+h) - f(x)}{h} - d + d' - \frac{f(x+h) - f(x)}{h}\right\|_Y \leqslant \left\|\frac{f(x+h) - f(x)}{h} - d\right\|_Y + \left\|\frac{f(x+h) - f(x)}{h} - d'\right\|_Y < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon = \|d' - d\|_Y \Rightarrow \|d - d'\|_Y < \|d - d'\|_Y$ a contradiction so we must have $d' = d$.
2. Assume that there exists two different left derivatives d, d' at x define then $\varepsilon = \|d' - d\|_Y > 0$, then there exists $\delta, \delta' > 0$ such that if $0 < h < \delta$ and $x - h \in [a, b]$ that $\left\|\frac{f(x-h) - f(x)}{f} - d\right\|_Y < \frac{\varepsilon}{2}$ and if $0 < h < \delta'$ and $x - h \in [a, b]$ then $\left\|\frac{f(x-h) - f(x)}{h} - d'\right\|_Y < \frac{\varepsilon}{2}$. As $x \in]a, b]$ we have $a < x \leqslant b$ there exists a $\delta'' > 0$ such that $a < x - \delta'' \leqslant b$ take then $h = \min(\delta'', \delta', \delta)$ then if $0 < h < \delta, \delta'$ then $x - h \in [a, b]$ and $\|d' - d\|_Y < \left\|\frac{f(x-h) - f(x)}{h} - d + d' - \frac{f(x-h) - f(x)}{h}\right\|_Y \leqslant \left\|\frac{f(x-h) - f(x)}{h} - d\right\|_Y + \left\|\frac{f(x-h) - f(x)}{h} - d'\right\|_Y < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon = \|d' - d\|_Y \Rightarrow \|d - d'\|_Y < \|d - d'\|_Y$ a contradiction so we must have $d' = d$.
3. Let $d = f'_+(x) = f'_-(x)$ and let $\varepsilon > 0$ then there exists $\delta', \delta'' > 0$ such that if $0 < h < \delta'$ and $x + h \in [a, b]$ then $\left\|\frac{f(x+h) - f(x)}{h} - d\right\|_Y < \varepsilon$ and if $0 < h < \delta''$ and $x - h \in [a, b]$ then $\left\|\frac{f(x-h) - f(x)}{-h} - d\right\|_Y < \varepsilon$. Take then $0 < |h| < \min(\delta', \delta'')$ such that $x + h \in]a, b[$ then two cases exists:

   a. $(\boldsymbol{h > 0})$ then as $0 < |h| < \min(\delta', \delta'') \Rightarrow 0 < h < \delta'$ and thus $\left\|\frac{f(x+h) - f(x)}{h} - d\right\|_Y < \varepsilon$

   b. $(\boldsymbol{h < 0})$ then as $0 < |h| < \min(\delta', \delta'') \Rightarrow 0 < |h| < \delta''$ and thus $\left\|\frac{f(x+h) - f(x)}{h} - d\right\|_Y \underset{h=-|h|}{=} \left\|\frac{f(x-|h|) - f(x)}{-|h|} - d\right\|_Y < \varepsilon$

   So we have in all cases that $\left\|\frac{f(x+h) - f(x)}{h} - d\right\|_Y < \varepsilon$ proving that $f'(x) = d = f'_-(x) = f'_+(x)$. On the other hand if $f'(x)$ exists then given $\varepsilon > 0$ there exists a $\delta' > 0$ such that if $0 < |h| < \delta$ and $x + h \in]a, b[$ then we have $\left\|\frac{f(x+h) - f(x)}{h} - f'(x)\right\| < \varepsilon$. Also as $x \in]a, b[$ we have $0 < x - a, b - x$ and there exists a δ'' such that $0 < \delta'' < \min(x - a, b - x)$ so if $0 < h < \delta''$ then $x - h > x - \delta'' >$

$x-(x-a)=a$ and $x+h<x+\delta''<x+b-x=b$ giving $x-h, x+h\in]a,b[$ so if $0<h<\delta=\min(\delta',\delta'')$ then we have

a. $x+h\in[a,b]$ and $\left\|\frac{f(x+h)-f(x)}{h}-f'(x)\right\|<\varepsilon$ proving that $f'_+(x)$ exists and is equal to $f'(x)$.

b. $x-h\in[a,b]$ and as $0<|-h|<\delta$ we have $\left\|\frac{f(x-h)-f(x)}{-h}-f'(x)\right\|<\varepsilon$ proving that $f'_-(x)$ exists and is equal to $f'(x)$

4. We must consider the following cases

   a. $(\boldsymbol{x}\in]\boldsymbol{a},\boldsymbol{b}[)$ then $\exists\delta'>0$ such that if $|h|<\delta'$ then $x+h\in]a,b[$, also as $(f^U)'$ exists on U and $[a,b]\subseteq U$ given $\varepsilon>0$ there exists a $\delta''>0$ so that if $0<|h|<\delta$ and $h\in U_x$ then $\left\|\frac{f^U(x+h)-f^U(x)}{h}-(f^U)'(x)\right\|_Y<\varepsilon$. So if $0<|h|<\delta=\min(\delta',\delta'')$ then $\left\|\frac{f(x+h)-f(x)}{h}-(f^U)'(x)\right\|_Y=\left\|\frac{f^U(x+h)-f^U(x)}{h}-(f^U)'(x)\right\|_Y<\varepsilon$ proving that $f'(x)$ exists and $(f^U)'(x)=f'(x)$

   b. $(\boldsymbol{x}=\boldsymbol{a})$ then as $(f^U)'$ exists on U and $[a,b]\subseteq U$ given $\varepsilon>0$ there exists a $\delta>0$ such that if $0<|h|<\delta$ and $h\in U_a$ then $\left\|\frac{f^U(a+h)-f^U(a)}{h}-(f^U)'(a)\right\|_U<\varepsilon$. So if $0<h<\delta$ and $a+h\in[a,b]\Rightarrow h\in U_a$ then we have $\left\|\frac{f^U(a+h)-f^U(a)}{h}-(f^U)'(a)\right\|<\varepsilon$ proving that $f'_+(a)$ exists and $f'_+(a)=(f^U)'(a)$.

   c. $(\boldsymbol{x}=\boldsymbol{b})$ then as $(f^U)'$ exists on U and $[a,b]\subseteq U$ given $\varepsilon>0$ there exists a $\delta>0$ such that if $0<|h|<\delta$ and $h\in U_a$ then $\left\|\frac{f^U(a+h)-f^U(a)}{h}-(f^U)'(a)\right\|_U<\varepsilon$. So if $0<h<\delta\Rightarrow 0<|-h|<\delta$ then if $b+(-h)\in[a,b]\Rightarrow -h\in U_b$ then we have $\left\|\frac{f(b-h)-f(b)}{-h}-(f^U)'(b)\right\|_Y=\left\|\frac{f^U(b+(-h))-f^U(b)}{-h}-(f^U)'(b)\right\|_Y<\varepsilon$ proving that $f'_-(b)$ exists and $(f^U)'(b)=f'_-(b)$. □

The above theorem leads to the following definition

Definition 14.83. *Let $\langle\mathbb{R},\|\rangle$ be the real space with the canonical norm, $\langle Y,\|\|_Y\rangle$ a normed space, $[a,b]\subseteq\mathbb{R}, a<b$ and $f:[a,b]\to Y$ and $x\in[a,b]$ then f has a derivative $f'(x)$ at x on $[a,b]$ if and only if either $x=a$ and $f'_+(a)$ exists, $x=b$ and $f'_-(b)$ exists or $x\in]a,b[$ and $f'_+(x)$, $f'_-(x)$ exists with $f'_+(x)=f'_-(x)$).*

Theorem 14.84. *Let $\langle\mathbb{R},\|\rangle$ be the real space with the canonical norm, $\langle Y,\|\|_Y\rangle$ a normed space, $[a,b]\subseteq\mathbb{R}, a<b$ then the following are equivalent for $f:[a,b]\to Y$*

1. *f is C^1 on $[a,b]$*
2. *We have the following*

   a. *$\forall t\in]a,b[$ we have the existence of the derivative $(f_{|]a,b[})'(t)$*

b. *There exists a right,left derivative* $f'_+(a)$, $f'_-(b)$ *at* a, b

c. f': $[a, b] \rightarrow Y$ *defined by* $x \rightarrow f'(x) = \begin{cases} f'_+(a) \text{ if } x = a \\ (f_{|]a,b[})'(x) \text{ if } x \neq a, b \\ f'_-(b) \text{ if } x = b \end{cases}$ *is continuous on* $[a, b]$

Further if f *is* C^1 *then* $\forall U$ *open,* f^U: $U \rightarrow Y$ *with* $[a, b] \subseteq U$ *and* $(f^U)_{|[a,b]} = f$ *we have that* $((f^U)')_{|[a,b]} = f'$

Proof. The proof is simple but elaborated

1. **($1 \Rightarrow 2$)** By definition we have a open set $U \subseteq \mathbb{R}$ with $[a, b] \subseteq U$ such that f^U: $U \rightarrow Y$ is C^1 and $(f^U)_{|[a,b]} = f$. Then using 14.49 we have that the derivative $(f^U)'$ is defined on every $x \in U$, is continuous at every $x \in U$ and $D^1(f^U)(x)(1) = (f^U)'(x)$. We proceed now to prove (a),(b) and (c):

   a. If $x \in]a, b[\subseteq U$ then given $\varepsilon > 0$ there exists a $\delta > 0$ such that $\forall h \in U_x \vdash 0 < |h| < \delta$ we have $\left\| \frac{f^U(x+h) - f(x)}{h} - (f^U)'(x) \right\|_Y < \varepsilon$, but then for $h \in (]a, b[)_x \subseteq U_x$ we have $\left\| \frac{f_{|]a,b[}(x+h) - f_{|]a,b[}(x)}{h} - (f^U)'(x) \right\|_Y = \left\| \frac{f(x+h) - f(x)}{h} - (f^U)'(x) \right\|_Y = \left\| \frac{f^{(U)}(x+h) - f^U(x)}{h} - (f^U)'(x) \right\|_Y < \varepsilon$ proving that $(f_{|]a,b[})'(x)$ exists and $(f_{|]a,b[})'(x) = (f^U)'(x)$

   b. Let $\varepsilon > 0$ then

      i. As $a \in U$ there exists a δ such that $\forall h \in U_x \vdash 0 < |h| < \delta$ we have $\left\| \frac{f^U(a+h) - f(a)}{h} - (f^U)'(a) \right\|_Y < \varepsilon \Rightarrow$ if $0 < h < \delta$ and $a + h \in [a, b] \Rightarrow h \in [a, b]_a \subseteq U_a \Rightarrow 0 < |h| < \delta$ and $\left\| \frac{f(a+h) - f(a)}{h} - (f^U)'(a) \right\|_Y = \left\| \frac{f^U(a+h) - f^U(a)}{h} - (f^U)'(a) \right\| < \varepsilon$ proving that $f'_+(a)$ exists and $f'_+(a) = (f^U)'(a)$ (as uniqueness is garanteed by the previous lemma)

      ii. As $b \in U$ there exists a δ such that $\forall h \in U_x \vdash 0 < |h| < \delta$ we have $\left\| \frac{f^U(b+h) - f(b)}{h} - (f^U)'(b) \right\|_Y < \varepsilon \Rightarrow$ if $0 < h < \delta$ and $b - h \in [a, b] \Rightarrow h \in [a, b]_{-h} \subseteq U_{-h} \Rightarrow 0 < |-h| < \delta$ and $-h \in U_b$ and $\left\| \frac{f(b-h) - f(b)}{-h} - (f^U)'(b) \right\|_Y = \left\| \frac{f^U(b-h) - f^U(b)}{-h} - (f^U)'(b) \right\| < \varepsilon$ proving that $f'_-(b)$ exists and $f'_-(b) = (f^U)'(b)$ (as uniqueness is garanteed by the previous lemma)

   c. So we have proved (a),(b) but also that $f' = ((f^U)')_{|[a,b]}$ which is continuous by 12.133.

2. **($2 \Rightarrow 1$)** Given f: $[a, b] \rightarrow Y$ satisfying (a) and (b) then given $\xi > 0$ we have $[a, b] \subseteq]a - \xi, b + \xi[$ a open set, define then $f^{]a-\xi,b+\xi[}$: $]a - \xi, b + \xi[\rightarrow Y$ defined by $x \rightarrow f^{]a-1,b+1[}(x) = \begin{cases} f(x) \text{ if } x \in [a, b] \\ f(a) + f'_+(a)(x-a) \text{if } x \in]a - \xi, a[\\ f(b) + f'_-(b)(x-b) \text{ if } x \in]b, b + \xi[\end{cases}$ then clearly we have that

$$(f^{]a-\xi,b+\xi[})_{|[a,b]} = f \tag{14.46}$$

Next we prove that $f^{]a-\xi,b+\xi[}$ has a derivative at every $x \in]a-\xi, b+\xi[$. The following cases occurs:

a. $(\boldsymbol{x}=\boldsymbol{a})$ Given $\varepsilon > 0$ there exists a $\delta > 0$ such that if $0 < h < \delta$ and $a+h \in [a,b]$ then $\left\|\frac{f(a+h)-f(a)}{h} - f'_+(a)\right\|_Y < \varepsilon$. Take then $\delta' = \min(\delta, b-a) > 0$ [as $a < b$]. Then if $0 < |h| < \delta'$ and $h \in]a-\xi, b+\xi[_a$ we have two subcases to consider

   i. $(\boldsymbol{h} < \mathbf{0})$ then $a+h \in]a-\xi, a[$ so that $\left\|\frac{f^{]a-\xi,b+\xi[}(a+h) - f^{]a-\xi,b+\xi[}(a)}{h} - f'_+(a)\right\|_Y = \left\|\frac{f(a)+f'_+(a)\cdot((a+h)-a)-f(a)}{h} - f'_+(a)\right\|_Y = \left\|\frac{f'_+(a)\cdot h}{h} - f'_+(a)\right\|_Y = 0 < \varepsilon$

   ii. $(\mathbf{0} < \boldsymbol{h})$ then $0 < h < \delta' \leqslant b-a \Rightarrow a < a+h \leqslant b \Rightarrow a+h \in [a,b] \underset{0<h<\delta}{\Rightarrow} \left\|\frac{f^{]a-\xi,b+\xi[}(a+h) - f^{]a-\xi,b+\xi[}(a)}{h} - f'_+(a)\right\|_Y = \left\|\frac{f(a+h)-f(a)}{h} - f'_+(a)\right\|_Y < \varepsilon$.

   So we have proved that $(f^{]a-\xi,b+\xi[})'(a)$ existsts and $(f^{]a-\xi,b+\xi[})'(a) = f'_+(a)$

b. $(\boldsymbol{x}=\boldsymbol{b})$ Given $\varepsilon > 0$ there exists a $\delta > 0$ such that if $0 < h < \delta$ and $b-h \in [a,b]$ then $\left\|\frac{f(b-h)-f(b)}{-h} - f'_-(b)\right\|_Y < \varepsilon$. Take then $\delta' = \min(\delta, b-a)$ then if $0 < |h| < \delta'$ and $h \in]a-\xi, b+\xi[_b$ we have two subcases to consider

   i. $(\boldsymbol{h} < \mathbf{0})$ then $0 < -h = |h| < \delta', b-a, \delta \Rightarrow a-h \leqslant b \Rightarrow a \leqslant b+h = b-|h| < b \Rightarrow b+h = b-|h| \in [a,b]$ so that $\left\|\frac{f^{]a-\xi,b+\xi[}(b+h) - f^{]a-\xi,b+\xi[}(b)}{h} - f'_-(b)\right\|_Y = \left\|\frac{f(b+h)-f(b)}{h} - f'_-(b)\right\|_Y = \left\|\frac{f(b-|h|)-f(b)}{-|h|} - f'_-(b)\right\|_Y < \varepsilon$

   ii. $(\mathbf{0} < \boldsymbol{h})$ then $b+h \in]b, b+\xi[$ and $\left\|\frac{f^{]a-1,b+1[}(b+h) - f^{]a-1,b+1[}(b)}{h} - f'_-(b)\right\|_Y = \left\|\frac{f(b)+f'_-(b)\cdot((b+h)-b)-f(b)}{h} - f'_-(b)\right\|_Y = \left\|\frac{f'_-(b)\cdot h}{h} - f'_-(b)\right\|_Y = 0 < \varepsilon$

   So we have proved that $(f^{]a-\xi,b+\xi[})'(b)$ exists and $(f^{]a-\xi,b+\xi[})'(b) = f'_-(b)$

c. $(\boldsymbol{x} \in]\boldsymbol{a}-\boldsymbol{\xi}, \boldsymbol{a}[)$ Given $\varepsilon > 0$ take then $\delta = a - x > 0$ then for a h such that $0 < |h| < \delta$ and $x \in]a-\xi, b+\xi[_x$ we have $x+h \in]a-\xi, b+\xi[$ and $x+h \leqslant x+|h| < x+\delta = x+(a-x) = a \Rightarrow x+h \in]a-1, a[$, so $\left\|\frac{f^{]a-\xi,b+\xi[}(x+h) - f^{]a-\xi,b+\xi[}(x)}{h} - f'_a(a)\right\|_Y = \left\|\frac{f(a)+f'_+(a)\cdot(x+h-a)-(f(a)-f'_+(x)\cdot(x-a))}{h} - f'_+(a)\right\|_Y = \left\|\frac{f'_+(a)\cdot h}{h} - f'_+(a)\right\|_Y = 0 < \varepsilon$. So we have proved that $(f^{]a-\xi,b+\xi[})'(x)$ exists and is equal to $f'_+(a)$

d. $(\boldsymbol{x}\in]\boldsymbol{b},\boldsymbol{b}+\boldsymbol{\xi}[)$ Given $\varepsilon>0$ take then $\delta=x-b>0$ then if h is such that $0<|h|<\delta$ and $h\in]a-\xi,b+\xi[_x$ then $x+h\in]a-\xi,b+\xi[$ and $h\leqslant|h|<x-b\Rightarrow b<x+h\Rightarrow x+h\in]b,b+\xi[$ and $\left\|\frac{f^{]a-\xi,b+\xi[}(x+h)-f^{]a-\xi,b+\xi[}(x)}{h}-f'_-(b)\right\|_Y=\left\|\frac{f(b)+f'_-(b)\cdot((x+h)-b)-(f(b)+f'_-(b)\cdot(x-b))}{h}-f'_-(b)\right\|_Y=\left\|\frac{f'_-(b)\cdot h}{h}-f'_-(b)\right\|_Y=0<\varepsilon$. So we have proved that $(f^{]a-\xi,b+\xi[})'(x)$ exists and is equal to $f'_-(b)$.

e. $(\boldsymbol{x}\in]\boldsymbol{a},\boldsymbol{b}[)$ Given $\varepsilon>0$ then by the hypothesis there exists a $\delta'>0$ such that if $0<|h|<\delta'$ and $h\in]a,b[_x$ then $\left\|\frac{f(x+h)-f(x)}{h}-(f_{|]a,b[})'(x)\right\|_Y<\varepsilon$. Take $\delta=\min(x-a,b-x,\delta')$ then if $0<|h|<\delta$ and $h\in]a-\xi,b+\xi[_x$ we have $h<|h|<x-a\Rightarrow a<x+h$ and $x+h<x+|h|\leqslant x+b-x=b\Rightarrow x+h<b$ proves that $x+h\in]a,b[$ so that $\left\|\frac{f^{]a-\xi,b+\xi[}(x+h)-f^{]a-\xi,b+\xi[}(x)}{h}-(f_{|]a,b[})'(x)\right\|_Y=\left\|\frac{f(x+h)-f(x)}{h}-(f_{|]a,b[})'(x)\right\|_Y<\varepsilon$. So we have proved that $(f^{]a-\xi,b+\xi[})'(x)$ exists and is equal to $(f_{|]a,b[})'(x)$

From (a),(b),(c),(d),(e) we have then that $(f^{]a-\xi,b+\xi[})'(x)$ exists $\forall x\in]a-\xi,b+\xi[$ and is equal to $f'_+(a)$ on $]a-\xi,a]$, $f'_-(b)$ on $[b,b+\xi[$ and to $(f_{|]a,b[})'(x)$. Now we must prove that it is continuous at every $x\in]a-\xi,b+\xi[$. The following cases must be considered

a. $(\boldsymbol{x}=\boldsymbol{a})$ As f' is continuous on $[a,b]$ we have $\forall\varepsilon>0$ the existence of a $\delta>0$ such that if $y\in[a,b]$ and $|y-a|<\delta\Rightarrow\|f'(y)-f'(a)\|_Y<\varepsilon$. So lets take $\delta'=\min(\delta,\frac{b-a}{2})$ and take y such that $|a-y|<\delta'$ then we have the following cases

   i. $(\boldsymbol{y}\leqslant\boldsymbol{a})$ Then $\|f^{]a-\xi,b+\xi['}(y)-f^{]a-\xi,b+\xi['}(a)\|_Y=\|f'_+(a)-f'_+(a)\|_Y=0<\varepsilon$

   ii. $(\boldsymbol{a}<\boldsymbol{y})$ Then as $|a-y|<\delta'\leqslant\frac{b-a}{2}\Rightarrow y-a<\frac{b-a}{2}\Rightarrow y<\frac{b+a}{2}<b$ and then $\|f^{]a-\xi,b+\xi['}(y)-f^{]a-\xi,b+\xi['}(a)\|_Y=\|f'_{|]a-b[}(y)-f'_+(a)\|=\|f'(y)-f'(a)\|_Y<\varepsilon$

   This proves continuity at $x=a$

b. $(\boldsymbol{x}=\boldsymbol{b})$ As f' is continuous on [a,b] we have $\forall\varepsilon>0$ the existence of a $\delta>0$ such that if $y\in[a,b]$ and $|y-b|<\delta\Rightarrow\|f'(y)-f'(b)\|_Y<\varepsilon$. Take now $\delta'=\min(\delta,\frac{b-a}{2})$ and lets take y such that $|b-y|<\delta'$ then we have the following cases

   i. $(\boldsymbol{y}<\boldsymbol{b})$ Then as $|b-y|<\delta'=\frac{b-a}{2}\Rightarrow b-y<\frac{b-a}{2}\Rightarrow b+\frac{a-b}{2}-y<0\Rightarrow\frac{a+b}{2}<y\Rightarrow a<y$ so we have $\|f^{]a-\xi,b+\xi['}(y)-f^{]a-\xi,b+\xi['}(b)\|_Y=\|f'_{|]a,b[}(y)-f'_-(b)\|=\|f'(y)-f'(b)\|_Y<\varepsilon$

ii. $(\boldsymbol{b} \leqslant \boldsymbol{y})$ Then $\|f^{]a-\xi,b+\xi[}(y) - f^{]a-\xi,b+\xi[}(b)\|_Y = \|f'_-(b) - f'_-(b)\|_Y = 0 < \varepsilon$

proving continuity at $x = b$.

c. $(\boldsymbol{x} \in]\boldsymbol{a}-\mathbf{1}, \boldsymbol{a}[)$ Take then $\delta = \min(a-x, x-(a-1)) > 0$ so that if $|y-x| < \delta$ we have $-\delta < y-x < \delta \Rightarrow x-\delta < y < \delta + x \Rightarrow x-(x-(a-1)) < y < a \Rightarrow a-1 < y < a \Rightarrow \|f^{]a-\xi,b+\xi['}(y) - f^{]a-\xi,b+\xi['}(x)\|_Y = \|f'_+(a) - f'_+(a)\|_Y = 0 < \varepsilon$ proving continuity at x.

d. $(\boldsymbol{x} \in]\boldsymbol{b}, \boldsymbol{b}+\mathbf{1}[)$ Take then $\delta = \min(x-b, b+1-x) > 0$ so that if $|y-x| < \delta$ we have $-\delta < y-x < \delta \Rightarrow -\delta + x < y < x+\delta \Rightarrow -(x-b)+x < y < x+b+1-x \Rightarrow b < y < b+1 \Rightarrow \|f^{]a-\xi,b+\xi['}(y) - f^{]a-\xi,b+\xi['}(x)\|_Y = \|f'_-(b) - f'_-(b)\|_Y = 0 < \varepsilon$ proving continuity at x

e. $(\boldsymbol{x} \in]\boldsymbol{a}, \boldsymbol{b}[)$ Because of continuity of $f'_{|]a,b[}$ on $[a,b]$ we have the existence of a $\delta > 0$ such that if $y \in]a,b[$ and $|x-y| < \delta \Rightarrow \|f'_{|]a,b[}(x) - f'_{|]a,b[}(y)\|_Y < \varepsilon$. Then if we take $\delta' = \min(\delta, x-a, b-x)$ and let y such that $|x-y| < \delta'$ then on one hand $|y-x| < \delta$ and on the other hand we have $-\delta' < y-x < \delta' \Rightarrow -\delta' + x < y < \delta' + x \Rightarrow -(x-a)+x < y < b-x+x \Rightarrow a < y < b$ so that $\|f^{]a-\xi,b+\xi['}(x) - f^{]a-\xi,b+\xi['}(y)\|_Y = \|f'_{|]a,b[}(x) - f'_{|]a,b[}(y)\|_Y < \varepsilon$, proving continuity at x.

From (a),(b),(c),(d) and (e) we have then that $(f^{]a-\xi,b+\xi[})'$ is continuous on $]a-\xi, b+\xi[$ proving that $f^{]a-\xi,b+\xi[}$ is C^1.

Further if f is C^1 then there exists a U open with $[a,b] \subseteq U$, $f^U: U \to Y$ such that $(f^U)_{|[a,b]} = f$ and f^U is C^1. Using then the previous lemma and (b) we have proved finally the theorem. □

Example 14.85. Let $\langle \mathbb{R}, \|\| \rangle$ be the real normed space, $\langle Y, \|\|_Y \rangle$ a normed space over $\mathbb{R}$, $x, y \in Y$ then $\sigma: [a,b] \to Y$ defined by $\sigma(t) = x + t \cdot y$ is C^1 and $\sigma'(t) = y$ if $t \in]a,b[$ together with $\sigma'_+(a) = y$, $\sigma'_-(b) = y$

Proof. Consider the following cases

1. $(\boldsymbol{t} = \boldsymbol{a})$ Given $\varepsilon > 0$ take $\delta = 1$ then if $0 < h < \delta$ such that $a+h \in [a,b]$ we have $\left\|\frac{\sigma(a+h) - \sigma(a)}{h} - y\right\|_Y = \left\|\frac{x + a \cdot y + h \cdot y - (x + a \cdot y)}{h} - y\right\|_Y = \left\|\frac{h \cdot y}{h} - y\right\|_Y = 0 < \varepsilon$ proving that $\sigma'_+(a) = y$

2. $(\boldsymbol{t} = \boldsymbol{b})$ Given $\varepsilon > 0$ take $\delta = 1$ then if $0 < h < \delta$ such that $b-h \in [a,b]$ we have $\left\|\frac{\sigma(b-h) - \sigma(b)}{-h} - y\right\|_Y = \left\|\frac{x + b \cdot y - h \cdot y - (x + b \cdot y)}{-h} - y\right\|_Y = \left\|\frac{-h \cdot y}{-h} - y\right\|_Y = 0 < \varepsilon$ proving that $\sigma'_-(b) = y$

3. $(\boldsymbol{t} \in]\boldsymbol{a}, \boldsymbol{b}[)$ Given $\varepsilon > 0$ take $\delta = 1$ then if $0 < |h| < \delta$ such that $h \in]a,b[_t \Rightarrow t+h \in]a,b[$ we have $\left\|\frac{\sigma(b+h) - \sigma(b)}{h} - y\right\|_Y = \left\|\frac{x + b \cdot y + h \cdot y - (x + b \cdot y)}{h} - y\right\|_Y = \left\|\frac{h \cdot y}{h} - y\right\|_Y = 0 < \varepsilon$ proving that $\sigma'_-(b) = y$ proving that $f'(t) = y$ □

14.4 Intermediate value and mean value theorems

Definition 14.86. *Let X be a topological space, $U \subseteq X$ open $f: U \rightarrow \mathbb{R}$ then $x \in U$ is a **local weak minimum (maximum) of f** if there exists a $\delta > 0$ such that $B_\delta(x) \subseteq U$ and $\forall y \in B_\delta(x)$ we have $f(x) \leqslant f(y)$ [or $f(y) \leqslant f(x)$]. A local minimum or maximum is also called a **local extremum**.*

Theorem 14.87. *Let $\langle \mathbb{R}, \| \rangle$ be the real normed space, $U \subseteq \mathbb{R}$ and $f: U \rightarrow \mathbb{R}$ be a function $x \in U$ a local extremum then if $f'(x)$ exists we have $f'(x) = 0$.*

Proof. We have two cases to consider

1. **(local weak maximum)** So there exists a $\delta > 0$ such that $]x - \delta, x + \delta[\subseteq U$ and $\forall y \in]x - \delta, x + \delta[$ we have that $f(y) \leqslant f(x)$. Assume now that $f'(x) = 0$ then two cases must be checked:

   a. **($f'(x) > 0$)** take then $\varepsilon = f'(x)$ then $\exists \delta' > 0$ such that if $0 < |h| < \delta'$, $h \in U_x$ then $\left| \frac{f(x+h) - f(x)}{h} - f'(x) \right| < \varepsilon$, so take h such that $0 < h < \min(\delta, \delta'), h \in U_x$ then $\frac{f(x+h) - f(x)}{h} \leqslant 0 < f'(x) \Rightarrow \varepsilon > \left| \frac{f(x+h) - f(x)}{h} - f'(x) \right| = f'(x) - \frac{f(x+h) - f(x)}{h} \geqslant f'(x) = \varepsilon \Rightarrow \varepsilon > \varepsilon$ a contradiction.

   b. **($f'(x) < 0$)** take then $\varepsilon = -f'(x)$ then $\exists \delta' > 0$ such that if $0 < |h| < \delta'$, $h \in U_x$ then $\left| \frac{f(x+h) - f(x)}{h} - f'(x) \right| < \varepsilon$, so take h such that $0 < h < \min(\delta' \delta'), h \in U_x$ then $\frac{f(x-h) - f(x)}{-h} \geqslant 0 > f'(x) \Rightarrow \varepsilon > \left| \frac{f(x-h) - f(x)}{-h} - f'(x) \right| = \frac{f(x-h) - f(x)}{-h} - f'(x) \geqslant -f'(x) = \varepsilon \Rightarrow \varepsilon > \varepsilon$ a contradiction

   as (a) and (b) yields a contradiction we must have $f'(x) = 0$.

2. **(local weak minimum)** So there exists a $\delta > 0$ such that $]x - \delta, x + \delta[\subseteq U$ and $\forall y \in]x - \delta, x + \delta[$ we have that $f(x) \leqslant f(y)$. Assume now that $f'(x) = 0$ then two cases must be checked:

   a. **($f'(x) > 0$)** take then $\varepsilon = f'(x) > 0$ then $\exists \delta' > 0$ such that if $0 < |h| < \delta', h \in U_x$ then $\left| \frac{f(x+h) - f(x)}{h} - f'(x) \right| < \varepsilon$, so take h such that $0 < |h| < \min(\delta, \delta'), h \in U_x$ then $\frac{f(x-h) - f(x)}{-h} \leqslant 0 < f'(x) \Rightarrow \varepsilon > \left| \frac{f(x-h) - f(x)}{-h} - f'(x) \right| = f'(x) - \frac{f(x-h) - f(x)}{-h} \geqslant f'(x) = \varepsilon \Rightarrow \varepsilon > \varepsilon$ a contradiction.

   b. **($f'(x) < 0$)** take then $\varepsilon = -f'(x) > 0$ then $\exists \delta' > 0$ such that if $0 < |h| < \delta', h \in U_x$ then $\left| \frac{f(x+h) - f(x)}{h} - f'(x) \right| < \varepsilon$, so take h such that $0 < |h| < \min(\delta, \delta'), h \in U_x$ then $\frac{f(x+h) - f(x)}{h} \geqslant 0 > f'(x) \Rightarrow \varepsilon > \left| \frac{f(x+h) - f(x)}{h} - f'(x) \right| = \frac{f(x+h) - f(x)}{h} - f'(x) > -f'(x) = \varepsilon \Rightarrow \varepsilon > \varepsilon$ a contradiction.

   as (a) and (b) yields a contradiction we must have $f'(x) = 0$ □

Theorem 14.88. (Rolle's theorem) *Let $[a, b] \subseteq \mathbb{R}, a < b$ $\langle \mathbb{R}, \| \rangle$ (which is a Banach space by 12.348) and $f \in \mathcal{C}([a, b], \mathbb{R})$ such that $f(a) = f(b) = 0$, differentiable on $]a, b[$ then there exists a $\zeta \in]a, b[$ such that $0 = f'(\zeta)$*

Proof. Using 12.443 there exists a $c,d\in\mathbb{R}$ such that $f([a,b])=[c,d]$ then as $c\leqslant f(a)=f(b)=0\leqslant d=c\leqslant 0\leqslant d$ we have the following cases to consider:

1. **($c=d=0$)** then f is constant and $\forall\zeta\in]a,b[$ we have $f'(\zeta)=\mathrm{Df}(\zeta)(1)=0$
2. **($c<0$)** As $f(a)=f(b)=0$ and $f([a,b])=[c,d]$ there exists a $\zeta\in]a,b[$ such that $f(\zeta)=c$ and as ζ is trivially a local minimum we have using the previous theorem (14.87) that $f'(\zeta)=0$.
3. **($0<d$)** As $f(a)=f(b)=0$ and $f([a,b])=[c,d]$ there exists a $\zeta\in]a,b[$ such that $f(\zeta)=d$ and as ζ is trivially a a local maximum we have using the previous theorem (14.87) that $f'(\zeta)=0$. □

Theorem 14.89. (Lagrange's theorem) *Let $[a,b]\subseteq\mathbb{R}, a<b$ $\langle\mathbb{R},\|\rangle$ and $f\in\mathcal{C}([a,b],\mathbb{R})$ which have a derivative everywhere in $]a,b[$ then there exists a $\zeta\in]a,b[$ such that $f(b)-f(a)=f'(\zeta)\cdot(b-a)$*

Proof. Define the function $g:[a,b]\rightarrow\mathbb{R}$ by $g(x)=f(x)-\left(f(a)+\frac{f(b)-f(a)}{b-a}\cdot(x-a)\right)$ which is differentiable in $]a,b[$ and continuous on $[a,b]$ [because constant functions are C^∞, linear functions in a finite dimensional spaces are continuous (see 12.176) and thus C^∞ (see 14.61) and the sum of C^∞ functions is C^∞, the same reasoning can be used for continuous functions]. Now $g(a)=f(a)-\left(f(a)+\frac{f(b)-f(a)}{b-a}\cdot(a-a)\right)=f(a)-f(a)=0$ and also $g(b)=f(b)-\left(f(a)+\frac{f(b)-f(a)}{b-a}\cdot(b-a)\right)=f(b)-f(a)-f(b)+f(a)=0$, applying Rolle's theorem we have then the existance of a $\zeta\in[a,b]$ such that $0=g'(\zeta)=0\Rightarrow 0=g'(\zeta)=f'(\zeta)-\frac{f(b)-f(a)}{b-a}\Rightarrow f(b)-f(a)=f'(\zeta)$ □

Corollary 14.90. *Let $U\subseteq\mathbb{R}$ be a open set $f:U\rightarrow U$ such that f has a derivative $\forall x\in U$, $\emptyset\neq A\subseteq U$ such that $\forall x\in A$ we have $0<f'(x)$ (or $0\leqslant f'(x)$) then f is strictly increasing (or increasing) on A and if $\forall x\in A$ we have that $f'(x)<0$ (or $f'(x)\leqslant 0$) then f is strictly decreasing (or decreasing) on A*

Proof. First as $\forall x\in U$ $f'(x)$ exists we have by 14.17 that f is differentiable at x (see 14.17) and thus continuous at x (see 14.10). So if $x,y\in A$ with $x<y$ we can apply Lagrange's theorem (see 14.89) so that $f(y)-f(x)=f'(\varepsilon)\cdot(y-x)$ where $\varepsilon\in]x,y[$. If now $\forall x\in A$ $0<f'(x)$ then $f'(\varepsilon)>0$, $y-x>0$ so that $f(y)-f(x)>0\Rightarrow f(y)>f(x)$ proving that f is strictly increasing on A. If $\forall x\in A$ $0\leqslant f'(x)$ then $f'(\varepsilon)\geqslant 0$, $y-x>0$ so that $f(y)-f(x)\geqslant 0\Rightarrow f(y)>f(x)$ proving that f is increasing on A.

If now $\forall x\in A$ we have $f'(x)<0$ [or $f'(x)\leqslant 0$] then we have that $(-f)'(x)=-f'(x)>0$ [or $(-f)'(x)=-f'(x)\geqslant 0$] so that we have that $-f$ is strictly increasing (or increasing) proving that f is strictly decreasing (decreasing). □

Definition 14.91. *Let $U\subseteq\mathbb{R}$ be a open set, $f:U\rightarrow\mathbb{R}$ a function $a,b\in U$ with $a\leqslant b$ and $[a,b]\subseteq U$ then f is convex (or concave) on $[a,b]$ if $\forall x\in[a,b]$ we have $f(a)+\frac{f(b)-f(a)}{b-a}\cdot(x-a)\geqslant f(x)$ (or $f(a)+\frac{f(b)-f(a)}{b-a}\cdot(x-a)\leqslant f(x)$). Essentially f is convex (concave) if a straight line between $f(a)$ and $f(b)$ lies above (below) the curve $f:[a,b]\rightarrow\mathbb{R}$.*

We can easely rewrite the this definition as follows

Lemma 14.92. *Let $U \subseteq \mathbb{R}$ be a open set, $f: U \rightarrow \mathbb{R}$ a function $a, b \in U$ with $a \leqslant b$ then f is convex (or concave) on $[a, b]$ if $\forall x \in [a, b]$ we have $\frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} \geqslant f(x)$ (or $\frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} \leqslant f(x)$)*

Proof. this follows from the following

$$\begin{aligned} f(a) + \frac{f(b) - f(a)}{b-a} \cdot (x-a) &= \frac{f(a) \cdot (b-a) + (f(b) - f(a)) \cdot (x-a)}{b-a} \\ &= \frac{f(a) \cdot b - f(a) \cdot a + f(b) \cdot x - f(b) \cdot a - f(a) \cdot x + f(a) \cdot a}{b-a} \\ &= \frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} \end{aligned}$$

□

Lemma 14.93. *Let $U \subseteq \mathbb{R}$ be a open set, $f: U \rightarrow \mathbb{R}$ a function $a, b \in U$ with $a \leqslant b$* and *$[a, b] \subseteq U$ then f is convex (or concave) on $[a, b]$ if $\forall \alpha, \beta$ with $0 \leqslant \alpha, \beta$ and $\alpha + \beta = 1$ we have that $f(\alpha \cdot a + \beta \cdot b) \geqslant \alpha \cdot f(a) + \beta \cdot f(b)$ (or $f(\alpha \cdot a + \beta \cdot b) \leqslant \alpha \cdot f(a) + \beta \cdot f(b)$)*

Proof. To prove this we show first the following equivalences

$$x \in [a, b] \Leftrightarrow \exists \alpha, \beta \vdash 0 \leqslant \alpha, \beta \wedge \alpha + \beta = 1 \text{ such that } x = \alpha \cdot a + \beta \cdot \tag{14.47}$$

If $x \in [a, b]$ then $0 \leqslant \frac{x-a}{b-a} = \beta, \frac{b-x}{b-a} = \alpha$ and $\alpha + \beta = \beta + \alpha = \frac{x-a+b-x}{b-a} = 1$ and $\alpha \cdot a + \beta \cdot b = \frac{(b-x) \cdot a + (x-a) \cdot b}{b-a} = \frac{b \cdot a - x \cdot a + x \cdot b - a \cdot b}{b-a} = \frac{x \cdot (b-a)}{b-a} = x$. Further if $\exists \alpha, \beta \vdash 0 \leqslant \alpha, \beta \wedge \alpha + \beta = 1$ then $0 \leqslant \alpha, \beta \leqslant 1$ so that $\alpha \cdot a + \beta \cdot b = \alpha \cdot a + (1-\alpha) \cdot b \geqslant_{a<b \wedge 0 \leqslant (1-\alpha)} \alpha \cdot a + (1-\alpha) \cdot a$ and $\alpha \cdot a + \beta \cdot b = \alpha \cdot a + (1-\alpha) \cdot b \leqslant_{a<b \wedge 0 \leqslant \alpha} \alpha \cdot b + (1-\alpha) \cdot b = b$ proving that $\alpha \cdot a + \beta \cdot b \in [a, b]$

From 14.48 it follows that we can always rewrite $\frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a}$ (where $x \in [a, b]$) as follows (where $0 \leqslant \alpha, \beta \wedge \alpha + \beta = 1$)

$$\begin{aligned} \frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} &= \frac{f(a) \cdot (b - \alpha \cdot a - \beta \cdot b) + f(b) \cdot (\alpha \cdot a + \beta \cdot b - a)}{b-a} \\ &= \frac{f(a) \cdot (b \cdot (1-\beta) - \alpha \cdot a) + f(b)(\beta \cdot b - a \cdot (1-\alpha))}{b-a} \\ &= \frac{f(a) \cdot (b \cdot \alpha - \alpha \cdot a) + f(b)(\beta \cdot b - a \cdot \beta)}{b-a} \\ &= \frac{f(a) \cdot (b-a) \cdot \alpha + f(b)(b-a) \cdot \beta}{b-a} \\ &= \alpha \cdot f(a) + \beta \cdot f(b) \end{aligned}$$

Hence as $f(x) = f(\alpha \cdot a + \beta \cdot b)$ we have that f is convex (concave) if $\alpha \cdot f(a) + \beta \cdot f(b) \geqslant f(\alpha \cdot a + \beta \cdot b)$ (or $\alpha \cdot f(a) + \beta \cdot f(b) \leqslant f(\alpha \cdot a + \beta \cdot b)$). □

The next theorem shows the relation between second derivate and convex and concave functions.

Theorem 14.94. *Let $U\subseteq\mathbb{R}$ be a open set, $f:U\rightarrow\mathbb{R}$ a function $a,b\in U$ with $a\leqslant b$ and $[a,b]\subseteq U$ such that $\forall x\in U$ we have that $f''(x)$ exists and $\forall x\in]a,b[$ we have $f''(x)\geqslant 0$ (or $f''(x)\leqslant 0$) then f is convex (or concave) on $[a,b]$*

Proof. First if $x=a$ (or $x=b$) then $\frac{f(a)\cdot(b-x)+f(b)\cdot(x-a)}{b-a}=f(a)$ (or $f(b)$) which is equal to $f(a)$ (or $f(b)$) so we must prove the theorem only for $x\in]a,b[$). Let $x\in]a,b[$

$$\begin{aligned}\frac{(x-a)}{(b-a)}+\frac{(b-x)}{(b-a)} &= \frac{x-a+b-x}{b-a}\\ &= 1\end{aligned}$$

so by multiplying both sides by $f(x)$ we have $f(x)=\frac{x-a}{b-a}\cdot f(x)+\frac{b-x}{b-a}\cdot f(x)$ so that

$$\begin{aligned}\frac{f(a)\cdot(b-x)+f(b)\cdot(x-a)}{b-a}-f(x) &= \frac{b-x}{b-a}\cdot f(a)+\frac{x-a}{b-a}\cdot f(b)-\frac{x-a}{b-a}\cdot f(x)-\frac{b-x}{b-a}\cdot f(x)\\ &= \frac{(x-a)}{b-a}\cdot(f(b)-f(x))-\frac{b-x}{b-a}\cdot(f(x)-f(a))\\ &= \frac{(x-a)\cdot(b-x)}{b-a}\cdot\frac{f(b)-f(x)}{b-x}-\frac{(b-x)\cdot(x-a)}{b-a}\cdot\frac{f(x)-f(a)}{x-a}\\ &= \frac{(x-a)\cdot(b-x)}{b-a}\left[\frac{f(b)-f(x)}{b-x}-\frac{f(x)-f(a)}{x-a}\right]\end{aligned}$$

Now as $\forall x\in U$ $f''(x)$ exists we have that $f'(x)$ exists and that f is continuous at x so using Lagrange's theorem (see 14.89) there exists a $y_1\in]x,b[$ and $y_2\in]a,x[$ such that $\frac{f(b)-f(x)}{(b-x)}=f'(y_1)$ and $\frac{f(x)-f(a)}{(x-a)}=f'(y_2)$ which using the above givc

$$\frac{f(a)\cdot(b-x)+f(b)\cdot(x-a)}{b-a}-f(x)=\frac{(x-a)\cdot(b-x)}{b-a}\cdot(f'(y_1)-f'(y_2)) \tag{14.48}$$

As $\forall x\in U$ $f''(x)$ exists we have that $f'(x)$ exists and that f' is continuous , further from $y_1\in]x,b[$ and $y_2\in]a,x[$ we have that $y_2<y_1$. Using 14.89 there exists a $z\in]y_2,y_1[$ such that $\frac{f'(y_1)-f'(y_2)}{(y_1-y_2)}=(f')'(z)=f''(x)$ which by substituting in 14.48 gives $\frac{f(a)\cdot(b-x)+f(b)\cdot(x-a)}{b-a}-f(x)=\frac{(x-a)\cdot(b-x)\cdot(y_1-y_2)}{b-a}\cdot f''(x)=a\cdot f''(x)$ where $\frac{(x-a)\cdot(b-x)\cdot(y_1-y_2)}{b-a}>0$. Hence if $f''(x)\geqslant 0$ (or $f''(x)\leqslant 0$) we have that $\frac{f(a)\cdot(b-x)+f(b)\cdot(x-a)}{b-a}-f(x)\geqslant 0$ (or $\frac{f(a)\cdot(b-x)+f(b)\cdot(x-a)}{b-a}-f(x)\leqslant 0$) proving our theorem. $\square$

We can use convexity or concavity as a condition to see if a extremum is a local minimum or a local maximum on a interval.

Corollary 14.95. *Let $U \subseteq \mathbb{R}$ be a open set, $f: U \to \mathbb{R}$ such that $\forall x \in U$ $f''(x)$ exists and $\exists x_0 \in U$ such that $f'(x) = 0$ then we have*

1. *If there exists a open V such that $x_0 \in V \subseteq U$ and $\forall x \in V$ $f''(x) \geqslant 0$ then x_0 is a local weak minimum of f*
2. *If there exists a open V such that $x_0 \in V \subseteq U$ and $\forall x \in V$ $f''(x) \leqslant 0$ then x_0 is a local weak maximum of f*

Proof. First from $f'(x_0) = 0$ we have using 14.87 that there exists a $\delta_1 > 0$ such that $]x_0 - \delta_1, x_0 + \delta_1[\subseteq U$ and we have either

$$(a)\ \forall x \in]x_0 - \delta_1, x_0 + \delta_1[\quad f(x) \leqslant f(x_0) \text{ or (b) } \forall x \in]x_0 - \delta_1, x_0 + \delta_1[f(x_0) \leqslant f(x) \tag{14.49}$$

Consider now the different cases of the corollary

1. As V is open there exists a $\delta_2 > 0$ such that $]x_0 - \delta_2, x_0 + \delta_2[\subseteq V$. If we now take $\delta = \min(\delta_1, \delta_2)$ then we have using 14.94 that

   $$]x_0 - \delta, x + \delta[\subseteq V,]x_0 - \delta_1, x_0 + \delta_1[\text{ and } \forall x \in]x_0 - \delta, x + \delta[\text{ we have } f''(x) \geqslant 0 \tag{14.50}$$

   Assume now that $\exists y \in]x_0 - \delta, x_0 + \delta[$ such that $f(y) < f(x_0)$ then we can not have $y = x_0$ so we must consider two cases

   $y < x_0$. then as $]y, x_0[\subseteq]x_0 - \delta, x_0 + \delta[$ we have by 14.94 that f is convex on $[y, x_0]$. As $y < x_0$ there exists a x such that $y < x < x_0 \Rightarrow x \in]y, x_0[$ so that by convexitiy we have

   $$\frac{f(y) \cdot (x_0 - x) + f(x_0) \cdot (x - y)}{x_0 - y} \geqslant f(x) \tag{14.51}$$

   As $0 < x_0 - x$, $f(y) < f(x_0) \Rightarrow (x_0 - x) \cdot f(y) < (x_0 - x) \cdot f(x_0)$ so that $\frac{f(x_0) \cdot (x_0 - x) + f(x_0) \cdot (x - y)}{x_0 - y} > \frac{f(y) \cdot (x_0 - x) + f(x_0) \cdot (x - y)}{x_0 - y}$ which as $\frac{f(x_0) \cdot (x_0 - x) + f(x_0) \cdot (x - y)}{x_0 - y} = \frac{f(x_0) \cdot x_0 - f(x_0) \cdot x + f(x_0) \cdot x - f(x_0) \cdot y}{x_0 - y} = \frac{f(x_0) \cdot (x_0 - y)}{x_0 - y} = f(x_0)$ gives $f(x_0) > \frac{f(y) \cdot (x_0 - x) + f(x_0) \cdot (x - y)}{x_0 - y}$. Using 14.51 we have then $f(x_0) > f(x)$ and as $x \in]x_0 - \delta, x_0 + \delta[\subseteq]x_0 - \delta_1, x_0 + \delta_1[$ we must choose in 14.49 the case (a) so that x_0 is a local weak minimum of f.

   $x_0 < y$. then as $]x_0, y[\subseteq]x_0 - \delta, x_0 + \delta[$ we have by 14.94 that f is convex on $[x_0, y]$. As $x_0 < y$ there exists a x such that $x_0 < x < y \Rightarrow x \in]x_0, y[$ so that by convexitiy we have

   $$\frac{f(x_0) \cdot (y - x) + f(y) \cdot (x - x_0)}{y - x_0} \geqslant f(x) \tag{14.52}$$

As $0 < x - x_0$, $f(y) < f(x_0) \Rightarrow (x - x_0) \cdot f(y) < (x - x_0) \cdot f(x_0)$ so that $\frac{f(x_0) \cdot (y - x) + f(x_0) \cdot (x - x_0)}{y - x_0} > \frac{f(x_0) \cdot (y - x) + f(y) \cdot (x - x_0)}{y - x_0}$ which as $\frac{f(x_0) \cdot (y - x) + f(x_0) \cdot (x - x_0)}{y - x_0} = \frac{f(x_0) \cdot y - f(x_0) \cdot x + f(x_0) \cdot x - f(x_0) \cdot x_0}{y - x_0} = \frac{f(x_0) \cdot (y - x_0)}{y - x_0} = f(x_0)$ gives $f(x_0) > \frac{f(y) \cdot (x_0 - x) + f(x_0) \cdot (x - y)}{x_0 - y}$. Using 14.52 we have then $f(x_0) > f(x)$ and as $x \in]x_0 - \delta, x_0 + \delta[\subseteq]x_0 - \delta_1, x_0 + \delta_1[$ we must choose in 14.49 the case (a) so that x_0 is a local weak minimum of f.

2. As V is open there exists a $\delta_2 > 0$ such that $]x_0 - \delta_2, x_0 + \delta_2[\subseteq V$. If we now take $\delta = \min(\delta_1, \delta_2)$ then we have using 14.94 that

$$]x_0 - \delta, x + \delta[\subseteq V,]x_0 - \delta_1, x_0 + \delta_1[\text{ and } \forall x \in]x_0 - \delta, x + \delta[\text{ we have } f''(x) \geqslant 0 \tag{14.53}$$

Assume now that $\exists y \in]x_0 - \delta, x_0 + \delta[$ such that $f(y) > f(x_0)$ then we can not have $y = x_0$ so we must consider two cases

$y < x_0$. then as $]y, x_0[\subseteq]x_0 - \delta, x_0 + \delta[$ we have by 14.94 that f is concave on $[y, x_0]$. As $y < x_0$ there exists a x such that $y < x < x_0 \Rightarrow x \in]y, x_0[$ so that by concavity we have

$$\frac{f(y) \cdot (x_0 - x) + f(x_0) \cdot (x - y)}{x_0 - y} \leqslant f(x) \tag{14.54}$$

As $0 < x_0 - x$, $f(y) > f(x_0) \Rightarrow (x_0 - x) \cdot f(y) > (x_0 - x) \cdot f(x_0)$ so that $\frac{f(x_0) \cdot (x_0 - x) + f(x_0) \cdot (x - y)}{x_0 - y} < \frac{f(y) \cdot (x_0 - x) + f(x_0) \cdot (x - y)}{x_0 - y}$ which as $\frac{f(x_0) \cdot (x_0 - x) + f(x_0) \cdot (x - y)}{x_0 - y} = \frac{f(x_0) \cdot x_0 - f(x_0) \cdot x + f(x_0) \cdot x - f(x_0) \cdot y}{x_0 - y} = \frac{f(x_0) \cdot (x_0 - y)}{x_0 - y} = f(x_0)$ gives $f(x_0) < \frac{f(y) \cdot (x_0 - x) + f(x_0) \cdot (x - y)}{x_0 - y}$. Using 14.54 we have then $f(x_0) < f(x)$ and as $x \in]x_0 - \delta, x_0 + \delta[\subseteq]x_0 - \delta_1, x_0 + \delta_1[$ we must choose in 14.49 the case (b) so that x_0 is a local weak maximum of f.

$x_0 < y$. then as $]x_0, y[\subseteq]x_0 - \delta, x_0 + \delta[$ we have by 14.94 that f is concave on $[x_0, y]$. As $x_0 < y$ there exists a x such that $x_0 < x < y \Rightarrow x \in]x_0, y[$ so that by convexitiy we have

$$\frac{f(x_0) \cdot (y - x) + f(y) \cdot (x - x_0)}{y - x_0} \leqslant f(x) \tag{14.55}$$

As $0 < x - x_0$, $f(y) > f(x_0) \Rightarrow (x - x_0) \cdot f(y) > (x - x_0) \cdot f(x_0)$ so that $\frac{f(x_0) \cdot (y - x) + f(x_0) \cdot (x - x_0)}{y - x_0} < \frac{f(x_0) \cdot (y - x) + f(y) \cdot (x - x_0)}{y - x_0}$ which as $\frac{f(x_0) \cdot (y - x) + f(x_0) \cdot (x - x_0)}{y - x_0} = \frac{f(x_0) \cdot y - f(x_0) \cdot x + f(x_0) \cdot x - f(x_0) \cdot x_0}{y - x_0} = \frac{f(x_0) \cdot (y - x_0)}{y - x_0} = f(x_0)$ gives $f(x_0) < \frac{f(y) \cdot (x_0 - x) + f(x_0) \cdot (x - y)}{x_0 - y}$. Using 14.55 we have then $f(x_0) < f(x)$ and as $x \in]x_0 - \delta, x_0 + \delta[\subseteq]x_0 - \delta_1, x_0 + \delta_1[$ we must choose in 14.49 the case (b) so that x_0 is a local weak maximum of f. □

Actually we can extend the above to maximum and minimums on a interval as follows

Corollary 14.96. *Let $U \subseteq \mathbb{R}$ be a open set, $f: U \to \mathbb{R}$ a function such that $\forall x \in U$ $f''(x)$ exists and $f''(x) \geqslant 0$ [or $f''(x) \leqslant 0$] then if $a, b \in U$ with $a \leqslant b$, $[a, b] \subseteq U$ such that $\exists x_0 \in [a, b]$ with $f'(x_0) = 0$ we have that $\forall x \in [a, b]$ $f(x) \geqslant f(x_0)$ (x_0 is a minimum of f on $[a, b]$) [or $\forall x \in [a, b]$ $f(x) \leqslant f(x_0)$ (x_0 is a maximum of f on $[a, b]$)]*

Proof. If $a = b \Rightarrow x[a, b] = \{x_0\}$ so that the theorem is trivial, so we may assume that $a < b$. From the previous corollary (see 14.95) we conclude that x_0 is a local weak minimum (local weak maximum) of f hence there exists a $\delta_1 > 0$ such that

$$\forall x \in]x_0 - \delta_1, x_0 + \delta_1[\subseteq U \text{ we have } f(x) \geqslant f(x_0) \, [\text{or } f(x) \leqslant f(x_0)] \tag{14.56}$$

Let now be $y \in [a, b]$ such that $f(y) < f(x_0)$ [or $f(y) > f(x_0)$] then we can not have $y = x_0$ so we must have either

***y* < *x*₀.** Then $\max(x_0 - \delta_1, y) < x_0$ and there exists a x such that $x_0 - \delta_1$, $y < z < x_0$. Applying 14.56 we have that

$$f(z) \geqslant f(x_0) \, [\text{or } f(z) \leqslant f(x_0)] \tag{14.57}$$

Using 14.94 and the hypothese we have that f is convex [or concave] on $[y, x_0]$. So we have that $\frac{f(y) \cdot (x_0 - z) + f(x_0) \cdot (z - y)}{x_0 - y} \geqslant f(z)$ [or $\frac{f(y) \cdot (x_0 - z) + f(x_0) \cdot (z - y)}{x_0 - y} \leqslant f(z)$]. Further from $0 < x_0 - z$ and $f(y) < f(x_0)$ [or $f(x_0) < f(y)$] so that $(x - x_0) \cdot f(y) < (x - x_0) \cdot f(x_0)$ [or $(x - x_0) \cdot f(y) > (x - x_0) \cdot f(x_0)$] we have that $f(x_0) = \frac{f(x_0) \cdot (x_0 - z) + f(x_0) \cdot (z - y)}{x_0 - y} > f(z)$ [$f(x_0) = \frac{f(x_0) \cdot (x_0 - z) + f(x_0) \cdot (z - y)}{x_0 - y} < f(z)$] and thus $f(x_0) > f(z)$ [or $f(x_0) < f(z)$]. Applying then 14.57 we reach the contradiction $f(x_0) > f(x_0)$ [or $f(x_0) < f(x_0)$]. So we must conclude that

$$\forall y \in [a, b] \text{ we have } f(y) \geqslant f(x_0) \, [\text{or } f(y) \leqslant f(x_0)]$$

***x*₀ < *y*.** Then $x_0 < \min(y, x_0 + \delta_1)$ and there exists a x such that $x_0 < z < y$, $x_0 + \delta_1$. Applying 14.56 we have that

$$f(z) \geqslant f(x_0) \, [\text{or } f(z) \leqslant f(x_0)] \tag{14.58}$$

Using 14.94 and the hypothese we have that f is convex [or concave] on $[x_0, y]$. So we have that $\frac{f(x_0) \cdot (y - z) + f(y) \cdot (z - x_0)}{y - x_0} \geqslant f(z)$ [or $\frac{f(x_0) \cdot (y - z) + f(y) \cdot (z - x_0)}{y - x_0} \leqslant f(z)$]. Further from $0 < z - x_0$ and $f(y) < f(x_0)$ [or $f(x_0) < f(y)$] so that $(z - x_0) \cdot f(y) < (z - x_0) \cdot f(x_0)$ [or $(z - x_0) \cdot f(y) > (z - x_0) \cdot f(x_0)$] we have that $f(x_0) = \frac{f(x_0) \cdot (y - z) + f(x_0) \cdot (z - x_0)}{y - x_0} > f(z)$ [$f(x_0) = \frac{f(x_0) \cdot (y - z) + f(x_0) \cdot (z - x_0)}{y - x_0} < f(z)$] and thus $f(x_0) > f(z)$ [or $f(x_0) < f(z)$]. Applying then 14.58 we reach the contradiction $f(x_0) > f(x_0)$ [or $f(x_0) < f(x_0)$]. So we must conclude that

$$\forall y \in [a, b] \text{ we have } f(y) \geqslant f(x_0) \, [\text{or } f(y) \leqslant f(x_0)]$$ □

Lemma 14.97. *Let $\langle X,\|\|\rangle$ be a Banach space, $\langle \mathbb{R},\|\rangle$ the real normed space, $[a,b]\subseteq\mathbb{R}, a<b$ $f:[a,b]\rightarrow X$ a continuous function then $F:[a,b]\rightarrow X$ where $F(x)=\int_a^x f$ is C^1 on $[a,b]$ and $F'=f$ (see 14.84 for a definition of $'$ on $[a,b]$)*

Proof. We have to take the following cases:

1. **($x=a$)** Given $\varepsilon>0$ then as f is continuous at a there exists a $\delta>0$ such that if $|y-a|<\delta\Rightarrow\|f(y)-f(a)\|<\varepsilon$ or if h is such that $0<h<\delta$ and $a+h\in[a,b]$ then if $y\in[a,a+h]$ we have $\|f(y)-f(a)\|<\varepsilon$. Also $\left\|\frac{F(a+h)-F(a)}{h}-f(a)\right\|=\left\|\frac{\int_a^{a+h}f-\int_a^a}{h}-f(a)\right\|=\left\|\frac{\int_a^{a+h}f}{h}-f(a)\right\|=\left\|\frac{\int_a^{a+h}f-h\cdot f(a)}{h}\right\|\underset{12.433}{=}\left\|\frac{\int_a^{a+h}f-\int_a^{a+1}f(a)}{h}\right\|\underset{12.432}{=}\left\|\frac{\int_a^{a+h}(f-f(a))}{h}\right\|=\frac{\|\int_a^{a+h}(f-f(a))\|}{|h|}\leqslant_{12.433}\frac{\int_a^h\|f-f(a)\|}{h}<_{12.433}\frac{\varepsilon\cdot h}{h}=\varepsilon$ proving that $F'_+(a)=f(a)$

2. **($x=b$)** Given $\varepsilon>0$ then as f is continuous at a there exists a $\delta>0$ such that if $|y-b|<\delta\Rightarrow\|f(y)-f(b)\|<\varepsilon$ or if h is such that $0<h<\delta$ and $b-h\in[a,b]$ then if $y\in[b-h,b]$ we have $\|f(y)-f(b)\|<\varepsilon$. Also $\left\|\frac{F(b-h)-F(b)}{-h}-f(b)\right\|=\left\|\frac{\int_a^{b-h}f-\int_a^b}{-h}-f(b)\right\|\underset{12.435}{=}\left\|\frac{-\int_{b-h}^b f}{-h}-f(b)\right\|=\left\|\frac{\int_{b-h}^b f-h\cdot f(b)}{h}\right\|=\frac{\|\int_{b-h}^b(f-f(b))\|}{h}\leqslant\frac{\int_{b-h}^b\|f-f(b)\|}{h}<\frac{\varepsilon\cdot h}{h}=\varepsilon$ proving that $F'_-(b)=f(b)$

3. **($x\in]a,b[$)** Given $\varepsilon>0$ then as f is continuous at a there exists a $\delta>0$ such that if $|y-b|<\delta\Rightarrow\|f(y)-f(b)\|<\varepsilon$. Now if h is such that $0<|h|<\delta$ and $x+h\in]a,b[$ then we have the following two cases to consider:

   a. **($h>0$)** If $y\in[x,x+h]$ we have $|y-x|<h<\delta\Rightarrow\|f(y)-f(x)\|<\varepsilon$. Also $\left\|\frac{F(x+h)-F(x)}{h}-f(x)\right\|=\left\|\frac{\int_a^{x+h}f-\int_a^x f}{h}-f(x)\right\|=\left\|\frac{\int_x^{x+h}f}{h}-f(x)\right\|=\left\|\frac{\int_x^{x+h}f-h\cdot f(x)}{h}\right\|=\frac{\|\int_x^{x+h}(f-f(x))\|}{h}\leqslant\frac{\int_x^{x+h}\|f-f(x)\|}{h}<\frac{\varepsilon\cdot h}{h}=\varepsilon$

   b. **($h<0$)** in this case $|h|=-h$ and if $y\in[x-|h|,x]$ we have $|y-x|<|h|<\delta\Rightarrow\|f(y)-f(x)\|<\varepsilon$. Also $\left\|\frac{F(x+h)-F(x)}{h}-f(x)\right\|=\left\|\frac{F(x-|h|)-F(x)}{-|h|}-f(x)\right\|=\left\|\frac{F(x-|h|)-F(x)+|h|f(x)}{-|h|}\right\|=\frac{\|F(x-|h|)-F(x)+|h|f(x)\|}{|h|}=\frac{\|\int_a^{x-|h|}f-\int_a^x f+|h|f(x)\|}{h}=\frac{\left\|-\int_{x-|h|}^x f+|h|\cdot f(x)\right\|}{h}=\frac{\left\|\int_{x-|h|}^x(f-f(x))\right\|}{h}\leqslant\frac{\int_{x-|h|}^x\|f-f(x)\|}{h}<\frac{h\cdot\varepsilon}{h}=\varepsilon$

   (a) and (b) proves then $F'(x)=f(x)$

Using 14.84 we have then that F' exists on $[a,b]$ and that F is C^1. □

We prove now the classical fundamental theorem of Calculus.

Theorem 14.98. (Fundamental Theorem of Calculus (Classical)) *Let $\langle Y, \|\|_Y\rangle$ be a Banach space, $\langle \mathbb{R}, \|\|\rangle$ the real normed space, $[a,b] \subseteq \mathbb{R}, a < b$ $f: [a,b] \to Y$ a C^1 function on $[a,b]$ then we have that $f(b) - f(a) = \int_a^b f'$*

Proof. First note that as f' is defined and continuous on $[a,b]$ (see 14.84) and that $\langle Y, \|\|\rangle$ is a Banach space we have that $\forall x \in [a,b]$ $\int_a^x f'$ is well defined (see 12.435). So the function $g: [a,b] \to Y$ defined by $x \to g(x) = f(a) + \int_a^x f'$ is well defined. Using the previous lemma we have then that g is C^1 and $g'(x) = (f(a) + \int_a^x f')' = 0 + f(x) = f'(x) \Rightarrow g' = f'$ proving that

$$g' = f' \tag{14.59}$$

Note also that $f(a) = g(a) + \int_a^a f' = g(a) + 0 = g(a)$ or

$$f(a) = g(a) \tag{14.60}$$

Consider now $\varphi:]a,b[\to \mathbb{R}$ defined by $x \to \varphi(x) = \|f(x) - g(x)\|$. Given $x \in]a,b[$ choose a $0 < \delta$ such that $B_{||}(x,\delta) \subseteq]a,b[$ then if $0 < |h| < \delta$ we have that $|x + h - x| = |h| < \delta \Rightarrow x + h \in B_{||}(x,\delta) \Rightarrow x + h \in]a,b[$ or

$$\text{If } 0 < |h| < \delta \Rightarrow x + h \in]a,b[\tag{14.61}$$

So we have that

$$\begin{aligned}
\left|\frac{\varphi(x+h) - \varphi(x)}{h}\right| &= \left|\frac{\|f(x+h) - g(x+h)\| - \|f(x) - g(x)\|}{h}\right| \\
&\leqslant \frac{\|f(x+h) - g(x+h) - f(x) + g(x)\|}{|h|} \\
&= \left\|\frac{(f(x+h)) - f(x) - (g(x+h) - g(x))}{h}\right\| \\
&\underset{f' = g' \text{ (see 14.59)}}{=} \left\|\frac{f(x+h) - f(x)}{h} - f'(x) - \left(\frac{g(x+h) - g(x)}{h} - g'(x)\right)\right\| \\
&\leqslant \left\|\frac{f(x+h) - f(x)}{h} - f'(x)\right\| + \left\|\frac{g(x+h) - g(x)}{h} - g'(x)\right\|
\end{aligned}$$

Now take $\varepsilon > 0$ then there exists a δ', δ'' such that if $0 < |h| < \delta', \delta''$ then $\left\|\frac{f(x+h) - f(x)}{h} - f'(x)\right\| < \frac{\varepsilon}{2}$ and $\left\|\frac{g(x+h) - g(x)}{h} - g'(x)\right\| < \frac{\varepsilon}{2}$ so if $0 < |h| < \delta''' = \min(\delta, \delta', \delta'')$ we have $\left\|\frac{\varphi(x+h) - \varphi(x)}{h} - 0\right\| < \varepsilon$ proving that

$$\forall x \in]a,b[\vDash \varphi'(x) = 0 \tag{14.62}$$

Given $x, y \in [a, b]$ with $x < y$ we have by Lagrange's Theorem (see 14.89) that there exists a $t \in]x, y[$ such that $\varphi(x) - \varphi(y) = \varphi'(t) \cdot (x - y) = 0 \cdot (x - y) = 0$ proving that φ is constant on $[a, b]$. As $\varphi(a) = \| f(a) - g(a)\| \underset{14.60}{=} 0$ we have that $\forall x \in [a, b] \vDash \varphi(x) = 0 \Rightarrow f(x) = g(x)$ so that $f(b) = g(b) = f(a) + \int_a^b f' \Rightarrow f(b) - f(a) = \int_a^b f'$ proving the theorem. □

Theorem 14.99. (Mean Value Theorem (first version)) *Let $\langle X, \|\|_X\rangle$ be a normed vector space, $U \subseteq X$, a open set, $\langle \mathbb{R}, \|\rangle$ and $f: U \rightarrow \mathbb{R}$ a function that is C^1 on U. If $x, y \in U$ such that $\forall t \in [0, 1]$ we have $(1 - t) \cdot x + t \cdot y = t \cdot y + (1 - t) \cdot x = x + t \cdot (y - x) \in U$ then there exists a $t_0 \in [0, 1]$ such that if $z = x + t_0 \cdot (x - y)$ we have $f(y) - f(x) = \mathrm{Df}(z)(y - z)$*

Proof. Define $\sigma: [0, 1] \rightarrow U$ by $t \rightarrow \sigma(t) = x + t \cdot (y - x)$ then $\sigma(0) = x$, $\sigma(1) = y$. Define then $g = f \circ \sigma: [0, 1] \rightarrow \mathbb{R}$, this function is continuous on $[0, 1]$ and differentiable on $]0, 1]$ (and thus on $]0, 1[$) [as σ is trivially C^∞ [sum of constant and linear function [that is continuous because of 12.296]]] and f is C^1 on U and $\sigma[0, 1] \subseteq U$]. By the Lagrange theorem (see 14.89) there exists a $\xi \in]0, 1[$ such that $g(1) - g(0) = g'(\xi) \cdot (1 - 0) = g'(\xi)$. Using the chain rule we have that $g'(\xi) = (D(g \circ f)(\xi))(1) = (\mathrm{Df}(\sigma(\xi)) \circ D\sigma(\xi))(1) = \mathrm{Df}(\sigma(\xi))(D\sigma(\xi)(1)) = \mathrm{Df}(\sigma(\xi))(\sigma'(\xi)) = \mathrm{Df}(\sigma(\xi))(y - x)$. Now $z = \sigma(\xi) = x + \xi \cdot (y - x)$ so that we have $f(y) - f(x) = f(\sigma(1)) - f(\sigma(0)) = g'(1) - g'(0) = \mathrm{Df}(z) \cdot (y - x)$ proving the theorem. □

Definition 14.100. *A subset C of a vector space X over $\mathbb{R}$ is convex if for every $t \in [0, 1]$ we have $\forall x, y \in C$ that $(1 - t) \cdot x + t \cdot y = x + t \cdot (y - x) \in C$*

Theorem 14.101. *Let $\langle X, \|\|\rangle$ be a normed space, $x \in X$ and $\delta > 0$ then $B_{\|\|}(x, \delta)$ is convex and $\bar{B}_{\|\|}(x, \delta)$ is convex.*

Proof. Let $t \in [0, 1]$ then for $x, y, z \in X$ we have

$$\begin{aligned}
\|y + t \cdot (z - y) - x\| &= \|y + t \cdot (z - y) - x + t \cdot x - t \cdot x\| \\
&= \|y \cdot 1 + t \cdot z - t \cdot y - x \cdot 1 + t \cdot x - t \cdot x\| \\
&= \|y \cdot (1 - t) - x \cdot (1 - t) + t \cdot z - t \cdot x\| \\
&= \|(y - x) \cdot (1 - t) + t \cdot (z - x)\| \\
&\leqslant (1 - t) \cdot \|y - x\| + t \cdot \|z - x\|
\end{aligned}$$

So if $y, z \in B_{\|\|}(x, \delta)$ then $\|y + t \cdot (z - y) - x\| \leqslant (1 - t) \cdot \|y - x\| + t \cdot \|z - x\| < (1 - t) \cdot \delta + t \cdot \delta = \delta \Rightarrow y + t \cdot (z - y) \in B_{\|\|}(x, \delta)$ also if $y, z \in \bar{B}_{\|\|}(x, \delta)$ then $\|y + t \cdot (z - y) - x\| \leqslant (1 - t) \cdot \|y - x\| + t \cdot \|z - x\| \leqslant (1 - t) \cdot \delta + t \cdot \delta = \delta \Rightarrow y + t \cdot (z - y) \in \bar{B}_{\|\|}(x, \delta)$ □

Theorem 14.102. (Mean Value Theorem (generalized)) *Let $\langle X, \|\|_X\rangle$ be a normed space over $\mathbb{R}$ and let $\langle Y, \|\|_Y\rangle$ be a Banach space over $\mathbb{R}$, $U \subseteq X$ a open set and $f: U \rightarrow Y$ a C^1 function. Assume that there is a convex subset $C \subseteq U$ and a constant $k \in \mathbb{R}$ such that $\forall x \in C$ we have $\|\mathrm{Df}(x)\| \leqslant k$ then $\forall x, y \in C$ we have $\|f(y) - f(x)\|_Y \leqslant k \cdot \|y - x\|_X$.*

Proof. Let $x,y\in C$ and define $\alpha:\mathbb{R}\rightarrow X$ by $t\rightarrow\alpha(t)=x+t\cdot(y-x)$ then by the fact that C is convex we have that $\alpha([0,1])\subseteq C\subseteq U$. Then as α is continuous [sum of constant function and a linear function that is continuous because of 12.296] we have that $V=\alpha^{-1}(U)$ is open. Define now $\sigma=\alpha_{|V}:V\rightarrow X$, then $\sigma(V)=\alpha(V)\subseteq U$ and because $\alpha([0,1])\subseteq U$ we have that $[0,1]\subseteq V$. Define now $g^V=f\circ\sigma:V\rightarrow Y$ then g^V is of class C^1 because of the Chain Rule and $\mathrm{Dg}^V(t)=\mathrm{Df}(\sigma(t))\circ D\sigma(t)$ and $(g^V)'(t)=\mathrm{Dg}^V(t)(1)=\mathrm{Df}(\sigma(t))(D\sigma(t)(1))=\mathrm{Df}(\sigma(t))(y-x)$. Define now h: $[0,1]\rightarrow X$ by $h=(g^V)_{|[0,1]}$ then using 14.84 we have that h' is defined on $[0,1]$ and $h'=((g^V)')_{|[0,1]}=((g^V)')_{|[0,1]}$. Next

$$\begin{aligned} f(y)-f(x) &= f(x+1\cdot(y-x))-f(x+0\cdot(y-x)) \\ &= (f\circ\sigma)(1)-(f\circ\sigma)(0) \\ &\underset{1,0\in[0,1]}{=} h(1)-h(0) \\ &\underset{14.98}{=} \int_0^1 h' \end{aligned}$$

so using 12.433 we have

$$\begin{aligned} \|f(y)-f(x)\|_Y &= \left\|\int_0^1 h'\right\|_Y \\ &\leqslant \int_0^1 \|h'\|_Y \end{aligned}$$

Finally if $t\in[0,1]$ we have $h'(t)=\mathrm{Df}(\sigma(t))(y-x)\Rightarrow\|h'(t)\|_Y\leqslant\|\mathrm{Df}(\sigma(t))\|\cdot\|y-x\|_X\leqslant k\cdot\|y-x\|_X$ and thus

$$\begin{aligned} \|f(y)-f(x)\|_Y &\leqslant \int_0^1 k\cdot\|y-x\|_X \\ &\leqslant (1-0)\cdot k\cdot\|y-x\|_X \\ &\leqslant k\cdot\|y-x\| \end{aligned}$$

proving the theorem. □

Definition 14.103. *Let X be a vector space, $a,b\in X$ then the line segment connecting a to b is the set $L_{a,b}=\{t\cdot a+(1-t)\cdot b|t\in[0,1]\}\subseteq X$. Note that $a,b\in L_{a,b}$ [take $t=0,1$]*

Example 14.104. $[0,1]=L_{0,1}$ in $\mathbb{R}$

Proof. If $x\in L_{0,1}$ then there exists a $t\in[0,1]$ such that $x=t\cdot 0+(1-t)\cdot 1=(1-t)\Rightarrow 0\leqslant 1-t\leqslant 1\Rightarrow L_{0,1}\subseteq[0,1]$. If $x\in[0,1]$ then $0\leqslant x\leqslant 1\Rightarrow 0\leqslant 1-x\leqslant 1\Rightarrow(1-x)\cdot 0+(1-(1-x))\cdot 1\in L_{0,1}\Rightarrow x\in L_{0,1}\Rightarrow[0,1]\subseteq L_{0,1}$. □

Lemma 14.105. *Let $\langle X,\|\|\rangle$ be a normed space then $a,b\in X$ then $L_{a,b}$ is convex, compact and thus bounded.*

Proof. If $x, y \in L_{a,b}$ then there exists $t_x, t_y \in [0,1]$ such that $x = t_x \cdot a + (1 - t_x) \cdot b$, $y = t_y \cdot a + (1 - t_y) \cdot b$. Take then $x + t \cdot (y - x) = t_x \cdot a + (1 - t_x) \cdot b + t \cdot (t_y \cdot a + (1 - t_y) \cdot b - t_x \cdot a - (1 - t_x) \cdot b) = t_x \cdot a + b - t_x \cdot b + t \cdot t_y \cdot a + t \cdot b - t \cdot t_y \cdot b - t \cdot t_x \cdot a - t \cdot b + t \cdot t_x \cdot b = t_x \cdot a + t \cdot t_y \cdot a - t \cdot t_x \cdot a + b - t_x \cdot b + t \cdot b - t \cdot t_y \cdot b - t \cdot b + t \cdot t_x \cdot b = (t_x + t \cdot t_y - t \cdot t_x) \cdot a + (1 - t_x + t - t \cdot t_y - t + t \cdot t_x) \cdot b = (t_x + t \cdot (t_y - t_x)) \cdot a + (1 - (t_x + t \cdot (t_y - t_x))) \cdot b = s \cdot a + (1 - s) \cdot b$ where $s = t_x + t \cdot (t_y - t_x) = t_x - t \cdot t_x + t \cdot t_y$. Now as $0 \leqslant t, t_x, t_y \leqslant 1$ we have that $t_x \cdot t \leqslant t_x \leqslant 1 \Rightarrow 0 \leqslant t_x - t_x \cdot t \Rightarrow 0 \leqslant t_x - t \cdot t_x + t \cdot t_y = t_x + t \cdot (t_y - t_x) \leqslant t_x + (t_y - t_x) = t_y \leqslant 1$ proving that $s \in [0,1]$ and thus that $x + t \cdot (y - x) \in L_{a,b}$.

Define $\sigma: [0,1] \rightarrow X$ by $t \rightarrow t \cdot a + (1 - t) \cdot b$ then σ is continuous for if $t \in [0,1]$, $\varepsilon > 0$ then take $\delta = \frac{\varepsilon}{\|a - b\|_Y + 1}$ and s such that $|s - t| < \delta$ then $\|\sigma(t) - \sigma(s)\|_Y = \|t \cdot a + (1 - t) \cdot b - s \cdot a - (1 - s) \cdot b\|_Y = \|(t - s) \cdot a + (s - t) \cdot b\|_Y = |t - s| \cdot \|a - b\|_Y < \delta \cdot \|a - b\| = \frac{e}{\|a - b\|_Y + 1} \cdot \|a - b\|_Y < \varepsilon$. Using 12.243 we have then that $L_{a,b} = \sigma([0,1])$ is compact, and by 12.242 we have then that $L_{a,b}$ is bounded. □

Corollary 14.106. *Let $\langle X, \|\|_X\rangle$ be a normed space over $\mathbb{R}$ and let $\langle Y, \|\|_Y\rangle$ be a Banach space over $\mathbb{R}$, $U \subseteq X$ a open set, $x, y \in U$ such that $L_{x,y} \subseteq U$ and $f: U \rightarrow Y$ a C^1 function then $\|f(x) - f(y)\|_Y \leqslant \sup(\{\|\mathrm{Df}(\xi)\| | \xi \in L_{x,y}\}) \cdot \|x - y\|_X$ [here $\|\|$ is the operator norm in $L(X, Y)$]*

Proof. As $L_{x,y}$ is compact by the previous lemma and Df is continuous [as f is C^1] we have that $\mathrm{Df}(L_{x,y})$ is compact (see 12.243) and thus by 12.242 bounded. So there exists a $M \geqslant 0$ such that $\forall L_1, L_2 \in \mathrm{Df}(L_{x,y})$ we have $\|L_1 - L_2\| \leqslant M$ so if $\xi \in L_{x,y}$ we have $\|\mathrm{Df}(\xi)\| = \|\mathrm{Df}(\xi) - \mathrm{Df}(x) + \mathrm{Df}(x)\| \leqslant \|\mathrm{Df}(\xi) - \mathrm{Df}(x)\| + \|\mathrm{Df}(x)\| \leqslant M + \|\mathrm{Df}(x)\|$ proving that $\{\|\mathrm{Df}(\xi)\| | \xi \in L_{x,y}\}$ is bounded above and thus by 9.43 that $\sup(\{\|\mathrm{Df}(\xi)\| | \xi \in L_{x,y}\})$ exists and is finite. Finally as $L_{x,y}$ is convex and $L_{x,y} \subseteq U$ we have by the Mean Value Theorem (14.102) that $\|f(x) - f(y)\|_Y \leqslant \sup(\{\|\mathrm{Df}(\xi)\| | \xi \in L_{x,y}\}) \cdot \|x - y\|_X$. □

Theorem 14.107. (Second Mean Value Theorem) *Let $\langle X, \|\|_X\rangle$ be a normed space over $\mathbb{R}$, $\langle Y, \|\|_Y\rangle$ a Banach space over $\mathbb{R}$. $a, b \in X$, $U \subseteq X$ a open set such that $L_{a,b} \subseteq U$ and $f: U \rightarrow Y$ a C^1 function then $\forall x \in U$ we have $\|f(a) - f(b) - \mathrm{Df}(x)(a - b)\|_Y \leqslant \|a - b\|_X \cdot \sup(\{\|\mathrm{Df}(\xi) - \mathrm{Df}(x)\| | \xi \in L_{a,b}\})$ (here $\|\|$ is the operator norm in $L(X, Y)$)*

Proof. Define $g: U \rightarrow Y$ by $\xi \rightarrow g(\xi) = f(\xi) - \mathrm{Df}(x)(\xi)$ which is C^1 as g is C^1 and $\mathrm{Df}(x)$ is C^∞ [as it is a continuous linear function]. Further if $\xi \in U$ then $\mathrm{Dg}(\xi) = \mathrm{Df}(\xi) - \mathrm{Df}(x)$ (as $\mathrm{Df}(x)$ is linear and continuous). Now using the Mean Value Theorem (see 14.106) we have $\|g(b) - g(a)\|_Y \leqslant \sup(\{\|\mathrm{Dg}(\xi)\| | \xi \in L_{a,b}\}) \cdot \|b - a\|_X \Rightarrow \|f(b) - \mathrm{Df}(x)(b) - f(a) + \mathrm{Df}(x)(a)\|_Y \leqslant \sup(\{\|\mathrm{Df}(\xi) - \mathrm{Df}(x)\| | \xi \in L\}) \Rightarrow \|f(b) - f(a) - \mathrm{Df}(x)(b - a)\|_Y \leqslant \sup(\{\|\mathrm{Df}(\xi) - \mathrm{Df}(x)\| | \xi \in L_{a,b}\})$. □

Another example of a theorem named Mean value theorem (see Dieudonne)

Theorem 14.108. *Let $[a, b] \subseteq \mathbb{R}$ be a non empty interval, $\langle Y, \|\|\rangle$ a Banach space, $f: I \rightarrow Y$, $\varphi: I \rightarrow \mathbb{R}$ continuous functions such that there exists a denumerable $E \subseteq [a, b]$ such that $\forall x \in [a, b] \setminus E$ f and φ have derivatives $f'(x)$ and $\varphi'(x)$ (see 14.83) with $\|f'(x)\| \leqslant \varphi'(x)$ then $\|f(b) - f(a)\| \leqslant \varphi(b) - \varphi(a)$*

Proof. As E is denumerable there exists a bijection $\sigma: \mathbb{N}_0 \to E$ Take now $\varepsilon > 0$ and define

$$\forall \beta \in [a, b] \text{ we define } \mathcal{E}(\beta) = \begin{cases} 0 \text{ if } \{i \in \mathbb{N}_0 | \rho(i) < \beta\} = \emptyset \\ \sum_{n \in \{i \in \mathbb{N}_0 | \rho(i) < \beta\}} \frac{1}{2^n} \end{cases} \tag{14.63}$$

and

$$A = \{\beta \in [a, b] | \forall \gamma \in [a, \beta[\vDash \| f(\gamma) - f(a) \| \leqslant \varphi(\gamma) - \varphi(a) + \varepsilon \cdot (\gamma - a) + \varepsilon \cdot \mathcal{E}(\gamma)\} \subseteq [a, b] \tag{14.64}$$

Then as $\| f(a) - f(a) \| = 0 = \varphi(a) - \varphi(a)$, $\varepsilon \cdot (a - a) = 0$ and $\mathcal{E}(a) = 0$ that

$$a \in A \Rightarrow A \neq \emptyset \tag{14.65}$$

Also if $\beta \in A$ then if $\xi \in [a, \beta[$ we have by the definition of A that also $\xi \in A$ proving together with $\beta \in A$ that

$$\forall \beta \in A \text{ we have } [a, \beta] \subseteq A \tag{14.66}$$

Take now $\sigma = \sup(A)$ then if $\xi \in [a, \sigma[$ there exists a $\beta \in A$ such that $a \leqslant \xi < \beta \leqslant \sigma \underset{14.66}{\Rightarrow} [a, \xi] \subseteq A \Rightarrow \xi \in A$ proving that $[a, \sigma[\subseteq A$. As $\beta \in A$ and $\xi \in [a, \beta[$ we have that $\| f(\xi) - f(a) \| \leqslant \varphi(\xi) - \varphi(a) + \varepsilon \cdot (\xi - a) + \varepsilon \cdot \mathcal{E}(\xi)$ proving as we have choosen $\xi \in [a, \sigma[$ that also $\sigma \in A$. So we have that

$$[a, \sigma] \subseteq A \tag{14.67}$$

If now $x \in A \setminus [a, \sigma]$ then as $A \subseteq [a, b]$ we have $a \leqslant x$ and thus must have $\sigma < x$ contradicting the fact that σ as a supremum is a upper bound of A so that $A \setminus [a, \sigma] = \emptyset$ proving finally that

$$[a, \sigma] = A \text{ where } \sigma = \sup(A) \tag{14.68}$$

Now if $\alpha, \beta \in [a, b]$ with $\alpha < \beta$ then if $n \in \{i \in \mathbb{N}_0 | \rho(i) < \alpha\}$ we have that $\rho(n) < \alpha < \beta \Rightarrow n \in \{i \in \mathbb{N}_0 | \rho(i) < \beta\}$ proving that $\{i \in \mathbb{N}_0 | \rho(i) < \alpha\} \subseteq \{i \in \mathbb{N}_0 | \rho(i) < \beta\}$ so that $\mathcal{E}(\alpha) \leqslant \mathcal{E}(\beta)$ proving

$$\forall \alpha, \beta \in [a, b] \text{ with } \alpha \leqslant \beta \text{ we have } \mathcal{E}(\alpha) \leqslant \mathcal{E}(\beta) \tag{14.69}$$

Now given $\xi > 0$, choosen arbitrary, then by the continuity of φ there exists a $\delta_\varphi > 0$ such that if $\beta \in [a, \sigma[\bigcap]\sigma - \delta_\varphi, \delta[$ then $|\sigma - \beta| < \delta_\varphi$ and $|\varphi(\beta) - \varphi(\sigma)| < \frac{\xi}{2} \Rightarrow \varphi(\beta) - \varphi(a) = \varphi(\sigma) - \varphi(a) + \varphi(\beta) - \varphi(\sigma) \leqslant \varphi(\sigma) - \varphi(a) + |\varphi(\beta) - \varphi(\sigma)| < \varphi(\sigma) - \varphi(a) + \frac{\xi}{2}$. By continuity of f there exists a $\delta_f > 0$ so that if $\beta \in [a, \sigma[\bigcap]\sigma - \delta_f, \sigma[$ then $\| f(\beta) - f(\sigma) \| < \frac{\xi}{2} \Rightarrow \| f(\sigma) - f(a) \| \leqslant \| f(\beta) - f(a) \| + \| f(\sigma) - f(b) \| < \| f(\beta) - f(a) \| + \frac{\xi}{2}$. So if $\beta \in [a, \sigma[\cap [\sigma - \min(\delta_\varphi, \delta_f), \sigma[$ then $\| f(\sigma) - f(a) \| \leqslant \| f(\beta) - f(a) \| + \frac{\xi}{2} \leqslant_{\text{definition of } A, \sigma \in A} \varphi(\beta) - \varphi(a) + \varepsilon \cdot (\beta - \alpha) + \varepsilon \cdot \mathcal{E}(\beta) + \frac{\xi}{2} < \varphi(\sigma) - \varphi(a) + \frac{\xi}{2} + \varepsilon \cdot \mathcal{E}(\beta) + \frac{\xi}{2} \leqslant_{\text{see } 14.69} \varphi(\sigma) - \varphi(a) + \varepsilon \cdot (\beta - \sigma) + \varepsilon \cdot \mathcal{E}(\sigma) + \xi \leqslant_{\beta < \sigma} \varphi(\sigma) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \mathcal{E}(\sigma)$. As $\xi > 0$ was choosen arbitrary we have using 9.56 that

$$\| f(\sigma) - f(a) \| \leqslant \varphi(\sigma) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \mathcal{E}(\sigma) \tag{14.70}$$

Suppose now that $\sigma < b$ then we have for σ to possibilities in relation to E

$\sigma \notin E$. then φ, f are differentiable at σ and as $\sigma < b$ we have by 14.83 that there exists $\lambda_\varphi, \lambda_f > 0$ such that for $\sigma \leqslant \xi < \lambda_\varphi < b$ we have $|\varphi(\xi) - \varphi(\sigma) - \varphi'(\sigma) \cdot (\xi - \sigma)| \leqslant \frac{\varepsilon}{2} \cdot (\xi - \sigma)$ and for $\sigma \leqslant \xi < \lambda_f < b$ we have $\|f(\xi) - f(\sigma) - f'(\sigma) \cdot (\xi - \sigma)\| \leqslant \frac{\varepsilon}{2} \cdot (\xi - \sigma)$ so that if $\sigma \leqslant \xi < \min(\lambda_\varphi, \lambda_f) < b$ we have

$$\begin{aligned}
\varphi'(\sigma) \cdot (\xi - \sigma) &= (\varphi'(\sigma) \cdot (\xi - \sigma) - (\varphi(\xi) - \varphi(\sigma)) + (\varphi(\xi) - \varphi(\sigma)) \\
&\leqslant |\varphi'(\sigma) \cdot (\xi - \sigma) - (\varphi(\xi) - \varphi(\sigma))| + \varphi(\xi) - \varphi(\sigma) \\
&= |\varphi(\xi) - \varphi(\sigma) - \varphi'(\sigma) \cdot (\xi - \sigma)| + \varphi(\xi) - \varphi(\sigma) \\
&\leqslant \varphi(\xi) - \varphi(\sigma) + \frac{\varepsilon}{2} \cdot (\xi - \sigma)
\end{aligned}$$

$$\begin{aligned}
\|f(\xi) - f(\sigma)\| &= \|f(\xi) - f(\sigma) - f'(\sigma) \cdot (\xi - \sigma) + f'(\sigma) \cdot (\xi - \sigma)\| \\
&\leqslant \|f(\xi) - f(\sigma) - f'(\sigma) \cdot (\mathcal{E} - \sigma)\| + \|f'(\sigma) \cdot (\xi - \sigma)\| \\
&\leqslant \frac{\varepsilon}{2} \cdot (\xi - \sigma) + \|f'(\sigma)\| \cdot (\xi - \sigma) \\
&\leqslant \frac{\varepsilon}{2} \cdot (\xi - \sigma) + \varphi'(\sigma) \cdot (\xi - \sigma) \text{ (by the hypothese of the theorem)} \\
&\leqslant \frac{\varepsilon}{2} \cdot (\xi - \sigma) + \varphi(\xi) - \varphi(\sigma) + \frac{\varepsilon}{2} \cdot (\xi - \sigma) \\
&= \varphi(\xi) - \varphi(\sigma) + \varepsilon \cdot (\xi - \sigma)
\end{aligned}$$

proving

$$\|f(\xi) - f(\sigma)\| \leqslant \varphi(\xi) - \varphi(\sigma) + \varepsilon \cdot (\xi - \sigma) \tag{14.71}$$

Next

$$\begin{aligned}
\|f(\xi) - f(a)\| &= \|f(\xi) - f(\sigma) + f(\sigma) - f(a)\| \\
&\leqslant \|f(\xi) - f(\sigma)\| + \|f(\sigma) - f(a)\| \\
&\leqslant \varphi(\xi) - \varphi(\sigma) + \varepsilon \cdot (\xi - \sigma) + \|f(\sigma) - f(a)\| \text{ (see 14.71)} \\
&\leqslant_{14.70} \varphi(\xi) - \varphi(\sigma) + \varepsilon \cdot (\xi - \sigma) + \varphi(\sigma) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \mathcal{E}(\sigma) \\
&= \varphi(\xi) - \varphi(a) + \varepsilon \cdot (\xi - a) + \varepsilon \cdot \mathcal{E}(\sigma) \\
&\leqslant_{14.69 \text{ and } \sigma < \xi} \varphi(\xi) - \varphi(a) + \varepsilon \cdot (\xi - a) + \varepsilon \cdot \mathcal{E}(\xi)
\end{aligned}$$

But this proves that $\min(\lambda_\varphi, \lambda_f) \in A$ wich is a contradiction as $\sup(A) = \sigma < \min(\lambda_\varphi, \lambda_f)$. So this case leads to a contradiction.

$\sigma \in E$. then there exists a $m \in \mathbb{N}_0$ such that $\sigma = \rho(m)$ then by the continuity of f, φ at σ together with $\sigma < b$ means that there exists μ_φ, μ_f such that if $\sigma < \xi < \mu_\varphi < b$ then $|\varphi(\xi) - \varphi(\sigma)| \leqslant \frac{\varepsilon}{2} \cdot \frac{1}{2^m}$ and if $\sigma < \xi < \mu_f < b$ then $\|f(\xi) - f(\sigma)\| < \frac{\varepsilon}{2} \cdot \frac{1}{2^m}$. So if $\sigma < \xi < \min(\mu_\varphi, \mu_f) < b$ then we have

$$\begin{aligned}
\|f(\xi) - f(a)\| &= \|f(\xi) - f(\sigma) + f(\sigma) - f(a)\| \\
&\leqslant \|f(\xi) - f(\sigma)\| + \|f(\sigma) - f(a)\| \\
&\leqslant \frac{\varepsilon}{2} \cdot \frac{1}{2^m} + \|f(\sigma) - f(a)\|
\end{aligned}$$

$$\begin{array}{rl}
\leqslant_{14.70} & \frac{\varepsilon}{2} \cdot \frac{1}{2^m} + \varphi(\sigma) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \mathcal{E}(\sigma) \\
= & \frac{\varepsilon}{2} \cdot \frac{1}{2^m} + \varphi(\xi) - \varphi(a) + \varphi(\sigma) - \varphi(\xi) + \varepsilon \cdot (\sigma - a) + \\
& \varepsilon \cdot \mathcal{E}(\sigma) \\
\leqslant & \frac{\varepsilon}{2} \cdot \frac{1}{2^m} + \varphi(\xi) - \varphi(a) + |\varphi(\sigma) - \varphi(\xi)| + \varepsilon \cdot (\sigma - a) + \\
& \varepsilon \cdot \mathcal{E}(\sigma) \\
\leqslant & \frac{\varepsilon}{2} \cdot \frac{1}{2^m} + \varphi(\xi) - \varphi(a) + \frac{\varepsilon}{2} \cdot \frac{1}{2^m} + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \mathcal{E}(\sigma) \\
= & \varphi(\xi) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \frac{1}{2^m} + \varepsilon \cdot \sum_{n \in \{i \in \mathbb{N} | \rho(i) < \sigma\}} \frac{1}{2^m} \\
\leqslant & \varphi(\xi) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \\
& \sum_{n \in \{i \in \mathbb{N} | \rho(i) < \xi\}} \frac{1}{2^m} (\text{as } \rho(m) = \sigma < \xi) \\
\leqslant_{\sigma < \xi} & \varphi(\xi) - \varphi(a) + \varepsilon \cdot (\xi - a) + \varepsilon \cdot \mathcal{E}(\xi)
\end{array}$$

proving again that $\min(\mu_\varphi, \mu_f) \in A$ wich is a contradiction as $\sup(A) = \sigma < \min(\lambda_\varphi, \lambda_f)$

As in all cases we have a contradiction we must have that $\sigma = b$ we have given $\varepsilon > 0$ arbitrary choosen that by 14.70 $\|f(b) - f(a)\| \leqslant \varphi(b) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \mathcal{E}(b)$ and thus by 9.56 that $\|f(b) - f(a)\| \leqslant \varphi(b) - \varphi(a)$ □

Corollary 14.109. *Let $[a, b] \subseteq \mathbb{R}$ be a non empty interval, $\langle Y, \|\|\rangle$ a Banach space, $f: [a, b] \rightarrow Y$ a coninuous function such that there exists a denumberable set $E \subseteq [a, b]$ so that $\forall x \in [a, b] \setminus E$ $f'(x)$ exists (see 14.83) and $\|f'(x)\| \leqslant M$ then $\|f(b) - f(a)\| \leqslant M \cdot (b - a)$*

Proof. Define $\varphi^{\mathbb{R}}: \mathbb{R} \rightarrow \mathbb{R}$ by $\varphi^{\mathbb{R}}(x) = M \cdot (x - a)$ then $\varphi'(x) = M$ so that by 14.82 $\varphi: [a, b] \rightarrow \mathbb{R}$ defined by $\varphi(x) = M \cdot (x - a)$ has $\forall x \in [a, b]$ the derivative $\varphi'(x) = M$. So $\forall x \in [a, b] \setminus E$ we have that $\|f'(x)\| \leqslant \varphi'(x)$ and thus by the previous theorem we have $\|f(b) - f(a)\| \leqslant \varphi(b) - \varphi(a) = M \cdot (b - a) - M \cdot (b - b) = M \cdot (b - a)$ □

Corollary 14.110. *Let $[a, b] \subseteq \mathbb{R}$ be a non empty interval, $\langle Y, \|\|\rangle$ a Banach space, $f: [a, b] \rightarrow Y$ a coninuous function so that $\forall x \in [a, b] \setminus E$ $f'(x)$ exists (see 14.83) and $\|f'(x)\| \leqslant M$ then $\|f(b) - f(a)\| \leqslant M \cdot (b - a)$*

Proof. As $\forall n \in \mathbb{N}_0$ we have that $0 < \frac{1}{n+1} < 1 \underset{0 < b - a}{\Rightarrow} 0 < (b - a) \cdot \frac{1}{n+1} < (b - a) \Rightarrow a < a + (b - a) \cdot \frac{1}{n+1} < b$ we have that $E = \left\{a + (b - a) \cdot \frac{1}{n+1} | n \in \mathbb{N}_0\right\} \subseteq [a, b]$ is a denumerable set and for $\forall x \in [a, b] \setminus E$ we have that $f'(x)$ exists (see 14.83) and $\|f'(x)\| \leqslant M$ so that by the previous corollary we have that $\|f(b) - f(a)\| \leqslant M \cdot (b - a)$ □

Theorem 14.111. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be Banach spaces, let $U \subseteq X$ a open set, $a, b \in U$ such that $L = \{a + (b - a) \cdot t | t \in [0, 1]\} \subseteq U$, $f: U \rightarrow Y$ a function such that $\forall x \in L$ we have that f is differentiable at x then*

$$\|f(b) - f(a)\|_Y \leqslant \|b - a\|_X \cdot \sup(\{\|Df(a + \lambda \cdot (b - a))\| | \lambda \in [0, 1]\})$$

Proof. Define the mapping $\varphi^{\mathbb{R}}: \mathbb{R} \rightarrow Y$ by $\varphi^{\mathbb{R}}(x) = a + (b - a) \cdot x$ which is C^{∞} (being a linear function+a constant function) take then the open set $V = \varphi^{\mathbb{R}-1}(U)$, if we then define $\varphi^V: V \rightarrow Y$ by $\varphi^V(x) = \varphi^{\mathbb{R}}(x)$ so that $\varphi^V([0,1]) = L_{a,b}$ so as f is differentiable on U and thus continuous we have that $f \circ \varphi^V: V \rightarrow Y$ is continuous and differentiable on every $x \in [0,1]$. Given $x \in [0,1]$ we have that $(f \circ \varphi^V)'(x) = D(f \circ \varphi^V)(x)(1) = (Df(\varphi^V(x)) \circ D\varphi^V(x))(1) = Df(a + (b-a) \cdot x)(D\varphi^V(x)(1)) = Df(a + (b-a) \cdot x)((b-a) \cdot 1) = Df(a + (b-a) \cdot x)(b-a)$ proving that if $f \circ \varphi$: $[0,1] \rightarrow Y$ defined by $(f \circ \varphi)(x) = (f \circ \varphi^V)(x)$ that

$$\forall x \in [0,1] \text{ we have } (f \circ \varphi)'(x) = (f \circ \varphi^V)'(x) = Df(a + (b-a) \cdot x) \cdot (b-a) \tag{14.72}$$

And thus $\forall x \in [0,1]$ we have $\|(f \circ \varphi)'(x)\|_Y = \|Df(a + (b-a) \cdot x)(b-e)\|_Y \leqslant \|Df(a + (b-a) \cdot x)\| \cdot \|b-a\|_X \leqslant \sup(\{\|Df(a + \lambda \cdot (b-a))\| | \lambda \in [0,1]\}) \cdot \|b-a\|_X$, so by the previous corollary we have then that $\|(f \circ \varphi)(1) - (f \circ \varphi)(0)\|_Y \leqslant \|b - a\|_X \cdot \sup(\{\|Df(a + \lambda \cdot (b-a))\| | \lambda \in [0,1]\}) \Rightarrow \|f(\varphi(1) - f(\varphi(0)))\|_Y \leqslant \|b-a\|_X \cdot \sup(\{\|Df(a + \lambda \cdot (b-a))\| | \lambda \in [0,1]\}) \Rightarrow \|f(b) - f(a)\|_Y \leqslant \|b-a\|_X \cdot \sup(\{\|Df(a + \lambda \cdot (b-a))\| | \lambda \in [0,1]\})$ □

14.5 Symmetry of Higher Order Differentials

Lemma 14.112. *Let $\langle X, \|\|_X \rangle$ be a normed space, $\langle Y, \|\|_Y \rangle$ be a Banach spaces, $U \subseteq X$ open and $f: U \rightarrow Y$ is C^2 then $\forall x \in U$ and $v, w \in X$ we have that $D^2 f(x)(v)(w) = D^2 f(x)(w)(v)$*

Proof. Let $x \in U$ then as U is open there exists a $r > 0$ so that $x \in B_{\|\|_X}(x, r) \subseteq U$, take then $v_1, v_2 \in X$ such that $\|v\|_X < \frac{r}{4}$, $\|w\|_X < \frac{r}{4}$ then $\forall t \in]-2, 2[\Rightarrow |t| < 2$ we have that $\|x + t \cdot v + w - x\|_X = |t| \cdot \|v\|_X + \|w\|_X \leqslant 2 \cdot \|v\|_X + \|w\|_X < 2 \cdot \frac{r}{4} + \frac{r}{4} < \frac{r}{2} + \frac{r}{2} = r \Rightarrow x + t \cdot v + w \in U$, $\|x + t \cdot v - x\|_X = \|t \cdot v\|_X = |t| \cdot \|v\|_X \leqslant 2 \cdot \|v\|_X < 2 \cdot \frac{r}{4} < r \Rightarrow x + t \cdot v \in U >$ So we can define the function $g:]-2, 2[\rightarrow Y$ by $t \rightarrow g(t) = f(x + t \cdot v + w) - f(x + t \cdot v)$, as $t \rightarrow x + t \cdot v + w$ and $t \rightarrow x + t \cdot v$ are trivially C^{∞} [being the sum of a constant and continuous linear function] we have that g is $C^2 \Rightarrow g$ is C^1. As $[0,1] = L_{0,1} \subseteq]-2, 2[$ we can use the Second Mean Value Theorem (see 14.107) to get $\|g(1) - g(0) - \mathrm{Dg}(0)(1-0)\|_Y \leqslant \sup(\{\|\mathrm{Dg}(\xi) - \mathrm{Dg}(0)\|\} | \xi \in [0, 1])$, now as $\mathrm{Dg}(0)(1) = g'(0)$ we have

$$\|g(1) - g(0) - g'(0)\|_Y \leqslant \sup(\{\|\mathrm{Dg}(\xi) - \mathrm{Dg}(0)\|\} | \xi \in [0,1]) \tag{14.73}$$

Now $\|g'(\xi) - g'(0)\|_Y = \|\mathrm{Dg}(\xi)(1) - \mathrm{Dg}(0)(1)\|_Y = \|(\mathrm{Dg}(\xi) - \mathrm{Dg}(0))(1)\|_Y \leqslant \|\mathrm{Dg}(\xi) - \mathrm{Dg}(0)\| \cdot |1| = \|\mathrm{Dg}(\xi) - \mathrm{Dg}(0)\|$ so that we have by 14.73 that

$$\|g(1) - g(0) - g'(0)\|_Y \leqslant \sup(\{\|g'(\xi) - g'(0)\|_Y | \xi \in [0,1]\}) \tag{14.74}$$

Using the Chain Rule and the differentiability of linear mappings we have that $g'(\xi) = \mathrm{Dg}(\xi)(1) = \mathrm{Df}(x + \xi \cdot v + w)(v \cdot 1) - \mathrm{Df}(x + \xi \cdot v)(v \cdot 1) = \mathrm{Df}(x + \xi \cdot v + w)(v) - \mathrm{Df}(x + \xi \cdot v)(v)$ giving

$$g'(\xi) = (\mathrm{Df}(x + \xi \cdot v + w) - \mathrm{Df}(x))(v) - (\mathrm{Df}(x + \xi \cdot v) - \mathrm{Df}(x))(v) \tag{14.75}$$

Since f is C^2 we have that Df is C^1 we can given a $\varepsilon > 0$ find a $\delta > 0$ such if $\|h\| < \delta$ and $h \in U_x$ then $\|\mathrm{Df}(x+h) - \mathrm{Df}(x) - D^2 f(x)(h)\|_Y < \varepsilon \cdot \|h\|_X$. Take now u, v such that $\|v\|_X, \|w\|_X \leqslant \min\left(\frac{r}{4}, \frac{\delta}{2}\right)$ then $\forall t \in [0,1]$ we have $\|t \cdot v + w\|_X \leqslant |t| \cdot \|v\|_X + \|w\|_X \leqslant 1 \cdot \frac{\delta}{2} + \frac{\delta}{2} = \delta$ and $x + t \cdot v + w \in U \Rightarrow t \cdot v + w \in U_x$, also $\|t \cdot v\|_X \leqslant \frac{\delta}{2} < \delta$ and $x + t \cdot v \in U \Rightarrow t \cdot v \in U_x$ so we have

$$\xi \in [0,1] \text{ then } \|\mathrm{Df}(x+\xi \cdot v + w) - \mathrm{Df}(x) - D^2 f(x)(\xi \cdot v + w)\|_Y \leqslant \varepsilon \cdot \|\xi \cdot v + w\|_X \tag{14.76}$$

$$\xi \in [0,1] \text{ then } \|\mathrm{Df}(x+\xi \cdot v) - \mathrm{Df}(x) - D^2 f(x)(\xi \cdot v)\|_Y \leqslant \varepsilon \cdot \|\xi \cdot v\|_X \tag{14.77}$$

Also

$$D^2 f(x)(\xi \cdot v + w)(v) - D^2 f(x)(\xi \cdot v)(v) \underset{D^2 f(x) \text{ is linear}}{=} D^2 f(x)(w)(v) \tag{14.78}$$

Now $\|g'(\xi) - D^2 f(x)(w)(v)\|_Y \underset{14.75}{=} \|(\mathrm{Df}(x+\xi \cdot v + w) - \mathrm{Df}(x))(v) - (\mathrm{Df}(x+\xi \cdot v) - \mathrm{Df}(x))(v) - D^2 f(x)(v)(w)\|_Y \underset{14.26}{=} \|(\mathrm{Df}(x+\xi \cdot v + w) - \mathrm{Df}(x) - D^2 f(x)(\varepsilon \cdot v + w))(v) - (\mathrm{Df}(x+\xi \cdot v) - \mathrm{Df}(x) - D^2 f(x)(\xi \cdot v))(v)\|_Y \leqslant \|(\mathrm{Df}(x+\xi \cdot v + w) - \mathrm{Df}(x) - D^2 f(x)(\varepsilon \cdot v + w))(v)\|_Y + \|(\mathrm{Df}(x+\xi \cdot v) - \mathrm{Df}(x) - D^2 f(x)(\xi \cdot v))(v)\|_Y \leqslant \|\mathrm{Df}(x+\xi \cdot v + w) - \mathrm{Df}(x) - D^2 f(x)(\varepsilon \cdot v + w)\| \cdot \|v\|_X + \|\mathrm{Df}(x+\xi \cdot v) - \mathrm{Df}(x) - D^2 f(x)(\xi \cdot v)\| \cdot \|v\|_X \leqslant \varepsilon \cdot \|\xi \cdot v + w\|_X \cdot \|v\| + \varepsilon \cdot \|\xi \cdot v\|_X \cdot \|v\| = \varepsilon \cdot \|v\| \cdot (\|\xi \cdot v + w\|_X + \|w\|_X) \leqslant \varepsilon \cdot \|v\|_X \cdot (|\xi| \cdot \|v\|_X + \|w\|_X + \|v\|_X) \leqslant \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X + \|v\|_X + \|w\|_X) = 2 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X)$ giving

$$\text{if } \xi \in [0,1] \text{ then } \|g'(\xi) - D^2 f(x)(w)(v)\|_Y \leqslant 2 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) \tag{14.79}$$

So if $t \in [0,1]$ we have as also $0 \in [0,1]$ that $\|g'(t) - g'(0)\|_Y \leqslant \|g'(t) - D^2 f(x)(w)(v) - (g'(0) - D^2 f(x)(w)(v))\|_Y \leqslant \|g'(t) - D^2 f(x)(w)(v)\|_Y + \|g'(0) - D^2 f(x)(w)(v)\|_Y \leqslant_{14.79} 4 \cdot \varepsilon \|v\|_X \cdot (\|v\|_X + \|w\|_X)$ giving

$$\|g'(t) - g'(0)\|_Y \leqslant 4 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) \tag{14.80}$$

Now the function $h: [-2,2] \to \mathbb{R}$ defined by $t \to h(t) = \|g'(t) - g'(0)\|_Y$ is continuous as g' is continuous [g is C^1 and 14.49]] and $\|\|$ is continuous (see 12.159). Using 12.292 there exists a $\xi_0 \in [0,1]$ so that

$$h(\xi_0) = \sup(h[[0,1]]) \Rightarrow \|g'(\xi_0) - g'(0)\|_Y = \sup(\{\|g'(\xi) - g'(0)\|_Y | \varepsilon \in [0,1]\}) \tag{14.81}$$

Then $\|g(1) - g(0) - D^2 f(x)(w)(v)\|_Y = \|g(1) - g(0) - g'(0) + g'(0) - g'(\xi_0) + g'(\xi_0) - D^2 f(x)(w)(v)\|_Y \leqslant \|g(1) - g(0) - g'(0)\|_Y + \|g'(0) - g'(\xi_0)\|_Y + \|g'(\xi_0) - D^2 f(x)(w)(v)\|_Y \leqslant_{14.74} \sup(\{\|g'(\xi) - g'(0)\| | \xi \in [0,1]\}) + \|g'(\xi_0) - g'(0)\|_Y + \|g'(\xi_0) - D^2 f(x)(w)(v)\|_Y \underset{14.81}{=} 2 \cdot \|g'(\xi_0) - g'(0)\|_Y + \|g'(\xi_0) - D^2 f(x)(w)(v)\|_Y \leqslant_{14.80,14.79} 4 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) + 2 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) = 6 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X)$ proving that

$$\|g(1) - g(0) - D^2 f(x)(w)(v)\|_Y \leqslant 6 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) \tag{14.82}$$

Now $g(1)-g(0)=f(x+v\cdot 1+w)-f(x+v\cdot 1)-(f(x+v\cdot 0+w)-f(x+v\cdot 0))=f(x+v+w)-f(x+v)-f(x+w)+f(x)$ is symmetric in v,w so if we take $v'=w$ and $w'=v$ with $\|v'\|_X,\|w'_Y\|\leqslant\min\left(\frac{r}{4},\frac{\delta}{2}\right)$ and $g'\colon[-2,2]\to Y$ by $g'(t)=f(x+t\cdot v'+w')-f(x+t\cdot v')$ then we have $\|g'(1)-g'(0)-D^2f(x)(w')(v')\|_Y<6\cdot\varepsilon\cdot\|v'\|_X\cdot(\|v'\|_X+\|w'\|_X)$ and as $g'(t)-g'(0)=g(t)-g(0)$ and $v'=w,w'=v$ we have

$$\|g(1)-g(0)-D^2f(x)(v)(w)\|_Y\leqslant 6\cdot\varepsilon\cdot\|w\|_X\cdot(\|w\|_X+\|v\|_X) \tag{14.83}$$

So $\|D^2f(x)(v)(w)-D^2f(x)(w)(v)\|_Y=\|g(1)-g(0)-D^2f(x)(w)(v)-(g(1)-g(0)-D^2f(x)(w)(v))\|_Y\leqslant\|g(1)-g(0)-D^2f(x)(w)(v)\|_Y+\|g(1)-g(0)-D^2f(x)(v)(w)\|_Y\leqslant 6\cdot\varepsilon\cdot\|v\|_x\cdot(\|v\|_X+\|w\|_X)+6\cdot\varepsilon\cdot\|w\|_X\cdot(\|w\|_X+\|v\|_X)=6\cdot\varepsilon\cdot(\|v\|_X+\|w\|_X)^2$ proving that

$$\|D^2f(x)(v)(w)-D^2f(x)(w)(v)\|_Y\leqslant 6\cdot\varepsilon\cdot(\|v\|_X+\|w\|_X)^2$$

Assume now that $\varepsilon'=\|D^2f(x)(v)(w)-D^2f(x)(w)(v)\|_Y>0$ then we can as ε in 14.83 was choosen arbitrary take $0<\varepsilon<\frac{\varepsilon'}{6\cdot(\|v\|_X+\|w\|_X)^2+1}$ then $\varepsilon'=\|D^2f(x)(v)(w)-D^2f(x)(w)(v)\|_Y<6\cdot\frac{\varepsilon'}{6\cdot(\|v\|_X+\|w\|_X)^2+1}\cdot(\|v\|_X+\|w\|_X)^2<\varepsilon'\Rightarrow\varepsilon'<\varepsilon'$ a contradiction so we conclude that $\forall u,v\in X$ with $\|u\|_X,\|v\|_X<\min\left(\frac{r}{4},\frac{\delta}{2}\right)$ we have $D^2f(x)(u)(v)=D^2f(x)(v)(w)$. Take now $u,v\in X$ and define $u'=\frac{u\cdot\min\left(\frac{r}{4},\frac{\delta}{2}\right)}{(\|u\|_X+1)}$, $v'=\frac{v\cdot\min\left(\frac{r}{4},\frac{\delta}{2}\right)}{(\|v\|_X+1)}$ then $\|u'\|_X=\frac{\|u\|_X}{(\|u\|_x+1)}\cdot\min\left(\frac{r}{4},\frac{\delta}{2}\right)<\min\left(\frac{r}{4},\frac{\delta}{2}\right)$, $\|v'\|_X=\frac{\|v\|_X}{(\|v\|_x+1)}\cdot\min\left(\frac{r}{4},\frac{\delta}{2}\right)<\min\left(\frac{r}{4},\frac{\delta}{2}\right)$ so that $D^2f(x)(v')(w')=D^2f(x)(w')(v')\underset{D^2f(x)\in L(X,L(X,Y))}{\Rightarrow}\frac{\min\left(\frac{r}{4},\frac{\delta}{2}\right)^2}{(\|u\|_X+1)\cdot(\|v\|_X+1)}\cdot D^2f(x)(v)(w)=\frac{\min\left(\frac{r}{4},\frac{\delta}{2}\right)^2}{(\|u\|_X+1)\cdot(\|v\|_X+1)}\cdot D^2f(x)(w)(v)\Rightarrow D^2f(x)(v)(w)=D^2f(x)(w)(v)$ proving at last our theorem. □

To prove that the n-th differential is symmetric we must first prove the following lemma and corollary.

Lemma 14.113. *Let $n\in\mathbb{N}$, X a set, Y a set and $f\colon X^{n+1}\to Y$ a function so that $\forall\sigma\in P_n$ and $\forall x=(x_1,\ldots,x_n)\in X^n$ and $x\in X$ we have $f(x_1,\ldots,x_n,x)=f(x_{\sigma(1)},\ldots,x_{\sigma(n)},x)$ and $\forall x=(x_1,\ldots,x_{n+1})\in X^{n+1}$ we have $f(x_1,\ldots,x_n,x_{n+1})=f(x_{1,\ldots,}x_{n-1},x_{n+1},x_n)$ then if $\tau\in P_{n+1}$ we have that $f(x_1,\ldots,x_{n+1})=f(x_{\tau(1)},\ldots,x_{\tau(n+1)})$*

Proof. The following cases exists

1. $(\boldsymbol{n=1})$ Then if $\sigma\in P_{n+1}$ we have either $\sigma=1_{[0,1]}$ and then $f(x_1,x_2)=f(x_{\sigma(1)},x_{\sigma(2)})$ or $\sigma=\left(1\leftrightarrow_2 2\right)$ and then $f(x_1,x_2)=f(x_2,x_1)=f(x_{\sigma(1)},x_{\sigma(2)})$ proving the theorem for the case $n=1$.
2. $(\boldsymbol{n>1})$ Let $\sigma\in P_{n+1}$ then we have the following cases to consider
   a. $(\boldsymbol{\sigma(n+1)=n+1})$ then $\tau=\sigma_{|\{1,\ldots,n\}}$ is a bijection so that $\tau\in P_n$ and thus $f(x_1,\ldots,x_n,x_{n+1})=f(x_{\tau(1)},\ldots,x_{\tau(n)},x_{n+1})=f(x_{\tau(1)},\ldots,x_{\tau(n)},x_{\sigma(n+1)})=f(x_{\sigma(1)},\ldots,x_{\sigma(n+1)})$

b. $\boldsymbol{(\sigma(n+1) \neq n+1)}$ then $\sigma(n+1) = k \in \{1, ..., n\}$, take then $i = \sigma^{-1}(n+1)$ so that $\sigma(i) = n+1$, as $\sigma(n+1) \neq n+1$ we have $i \in \{1, ..., n\}$. Then if $\tau = (i \leftrightarrow_n n)$ we have $f(x_{\sigma(1)}, ..., x_{\sigma(n)}, x_{\sigma(n+1)}) = f(x_{\sigma(\tau(1))}, ..., x_{\sigma(\tau(n))}, x_k) = f(x_{\sigma(\tau(1))}, ..., x_{\sigma(\tau(n-1))}, x_{\sigma(i)}, x_k) = f(x_{\sigma(\tau(1))}, ..., x_{\sigma(\tau(n-1))}, x_{n+1}, x_k) = f(x_{\sigma(\tau(1))}, ..., x_{\sigma(\tau(n-1))}, x_k, x_{n+1})$ so that

$$f(x_{\sigma(1)}, ..., x_{\sigma(n)}, x_{\sigma(n+1)}) = f(x_{\sigma(\tau(1))}, ..., x_{\sigma(\tau(n-1))}, x_k, x_{n+1}) \tag{14.84}$$

Now if $j \in \{1, ..., n-1\}$ then if $\sigma(\tau(j)) = n+1$ we have $\tau(j) = \sigma^{-1}(n+1) = i \Rightarrow (i \leftrightarrow_n n)(j) = i \Rightarrow j = n \notin \{1, ..., n-1\}$ a contradiction. So if we can define $\rho: \{1, ..., n\} \to \{1, ..., n\}$ definied by $j \to \rho(j) = \begin{cases} \sigma(\tau(j)) \text{ if } j \in \{1, ..., n-1\} \\ k \text{ if } j = n \end{cases}$. We prove next that ρ is a bijection and thus $\rho \in P_n$:

i. **(injectivity)** First if $k = \sigma(\tau(j)) \Rightarrow \sigma(n+1) = \sigma(\tau(j)) \Rightarrow n+1 = \tau(j)$ a contradiction (as $\tau \in P_n$) so $\forall j \in \{1, ..., n-1\}$ we have $\rho(j) \neq \rho(n)$. Assume now that $\exists j, l \in \{1, ...n-1\}$ such that $\rho(j) = \rho(l) \Rightarrow \sigma(\tau(j)) = \sigma(\tau(l)) \underset{\sigma \text{ is a bijection}}{\Rightarrow} \tau(j) = \tau(l) \underset{\tau \text{ is a bijection}}{\Rightarrow} j = l$. So $\forall j, l \in \{1, ..., \}$ we have $\rho(j) = \rho(l) \Rightarrow j = l$.

ii. **(surjectivity)** If $j \in \{1, ..., n\}$ then we have either

   A. $\boldsymbol{(j = k)}$ then $j = k = \rho(n)$

   B. $\boldsymbol{(j \in \{1, ..., n\} \setminus \{k\})}$ we have as σ is a bijection $r = \sigma^{-1}(j) \in \{1, ..., n+1\}$ and as $\sigma(n+1) = k$ we must have that $r \in \{1, ..., n\}$ and thus there exists a $s = \tau^{-1}(r)$. So that $\rho(s) = \sigma(\tau(s)) = \sigma(r) = j$

So using 14.84 we have then $f(x_{\sigma(1)}, ..., x_{\sigma(n+1)}) = f(x_{\rho(1)}, ..., x_{\rho(n)}, x_{n+1}) = f(x_1, ..., x_n, x_{n+1})$. □

Corollary 14.114. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X \rangle$ and $\langle Y, \|\|_Y \rangle$ normed sets and $L \in L^{n+1}(X; Y)$ is such that $\forall \sigma \in P_n$ and $\forall x_1, ..., x_n, x \in X$ that $L(x: x_1: ...: x_n) = L(x: x_{\sigma(1)}: ...: x_{\sigma(n)})$ and $\forall x_1, ..., x_{n-1}, x, y \in X$ that $L(x: y: x_1: ...: x_{n-1}) = L(y: x: x_1: ...: x_{n-1})$ then we have that $\forall \tau \in P_{n+1}$ and $\forall x_1, ..., x_{n+1} \in X$ that $L(x_1: ...: x_{n+1}) = L(x_{\sigma(1)}: ...: x_{\sigma(n+1)})$*

Proof. Define $S: L(X^{n+1}, Y) \to L(X^{n+1}, Y)$ by $L \to S(L)$ where $S(L)(x_1, ..., x_{n+1}) = L(x_{n+1}, ..., x_1)$ then define then $H: L^{n+1}(X; Y) \to L(X^{n+1}, Y)$ by $H = S \circ \mathcal{P}_{n+1}$ then we have

1. $\forall \sigma \in P_n$, $\forall x, x_1, ..., x_{n+1}$ that $H(L)(x_1, ..., x_n, x) = (S(\mathcal{P}_{n+1}(L)))(x_1, ..., x_n, x) = \mathcal{P}_{n+1}(L)(x, x_n, ..., x_1) = L(x: x_n: ...: x_1) \underset{10.67}{=} L(x: x_{\iota_n(1)}: ...: x_{\iota_n(n)}) = L(x: x_{\sigma(\iota_n(1))}: ...: x_{\sigma(\iota_n(n))}) = L(x: x_{\sigma(n)}: ...: x_{\sigma(1)}) = \mathcal{P}_{n+1}(x, x_{\sigma(n)}, ..., x_{\sigma(1)}) = (S(\mathcal{P}_{n+1}(L)))(x_{\sigma(1)}, ..., x_{\sigma(n)}, x) = H(x_1, ..., x_n, x)$

2. If $x_1, ..., x_{n+1} \in X$ then $H(L)(x_1, ..., x_n, x_{n+1}) = L(x_{n+1}: x_n: x_{n-1}: ...: x_1) = L(x_n: x_{n+1}: ...: x_1) = H(L)(x_1, ..., x_{n+1}, x_n)$

So we can use the previous lemma on $H(L)$ to find that given $\tau \in P_{n+1}$ then $L(x_1: \ldots: x_{n+1}) = H(L)(x_{n+1}, \ldots, x_1) \underset{10.67}{=} H(L)(x_{\iota_{n+1}(1)}, \ldots, x_{\iota_{n+1}(n+1)}) \underset{\text{previous lemma}}{=} H(L)(x_{\tau(\iota_{n+1}(1))}, \ldots, x_{\tau(\iota_{n+1}(n+1))}) = H(L)(x_{\tau(n+1)}, \ldots, x_{\tau(1)}) = L(x_{\tau(1)}: \ldots: x_{\tau(n+1)})$ □

Lemma 14.115. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ normed spaces and $m \in \{1, \ldots, n\}$ then if $(v_1, \ldots, v_m) \in X^m$ we define $e_{v_1, \ldots, v_m}: L^m(X; Y) \to Y$ by $L \to e_{v_1, \ldots, v_m}(L) = L(v_1: \ldots: v_m)$ then $e_{v_1, \ldots, v_m}$ is linear and continuous. So $e_{v_1, \ldots, v_m}$ is C^∞ and $\mathrm{De}_{v_1, \ldots, v_m}(L) = e_{v_1, \ldots, v_m}$ and if $n \geqslant 2$ then $D^n e_{v_1, \ldots, v_m}(L) = 0$ (the constant mapping to 0)*

Proof.

1. **(Linearity)** Then $e_{v_1, \ldots, v_m}(\alpha \cdot L_1 + \beta \cdot L_2) = (\alpha \cdot L_1 + \beta \cdot L_2)(v_1: \ldots: v_m) \underset{12.207}{=} \alpha \cdot L_1(v_1: \ldots: v_m) + \beta \cdot L_2(v_1: \ldots: v_m) = \alpha \cdot e_{v_1, \ldots, v_{n-2}}(L_1) + \beta \cdot e_{v_1, \ldots, v_{n-2}}(L_2)$
2. **(Continuity)** Then $\|e_{v_1, \ldots, v_{n-2}}(L)\|_Y = \|L(v_1: \ldots: v_{n-2})\|_Y \leqslant_{12.208} \|L\| \cdot \prod_{i \in \{1, \ldots, n-2\}} \|v_i\|_X$ proving that $\|e_{v_1, \ldots, v_n}\| \leqslant \|L\|$ and thus continuity.

The rest of the theorem follows from 14.61 □

Before we can prove the symmetry of the n-the differential we must first prove the following lemma's

Lemma 14.116. *Let $n \in \mathbb{N} \vdash n > 2$, $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be normed spaces, $U \subseteq X$ open, $v_1, \ldots, v_{n-2} \in X^{n-2}$, a $f: U \to Y$ that is C^n and $g: U \to Y$ defined by $x \to g(x) = D^{n-2} f(x)(v_1: \ldots v_{n-2})$ then g is C^2 and $D^2 g(x)(u)(v) = D^n f(x)(v: u: v_1: \ldots: v_{n-2})$*

Proof. As $g(x) = D^{n-2} f(x)(v_1: \ldots: v_{n-2}) = e_{v_1, \ldots, v_{n-2}}(D^{n-2} f(x)) = (e_{v_1, \ldots v_{n-2}} \circ D^{n-2} f)(x)$ proving that

$$g = (e_{v_1, \ldots, v_{n-2}} \circ D^{n-2} f) \tag{14.85}$$

so that by the previous lemma and the Chain Rule we have that g is C^2 and $\mathrm{Dg}(x) = \mathrm{De}_{v_1, \ldots, v_n}(D^{n-2} f(x)) \circ D((D^{n-2}) f)(x) = e_{v_1, \ldots, v_{n-2}} \circ D^{n-1} f(x)$ so that if $u \in X$ we have $\mathrm{Dg}(x)(u) = e_{v_1, \ldots, v_{n-2}}(D^{n-1} f(x)(u)) = (D^{n-1} f(x)(u))(v_1: \ldots: v_{n-2}) \underset{12.206}{=} D^{n-1} f(x)(u: v_1: \ldots: v_{n-2}) = e_{u, v_1, \ldots, v_{n-1}}(D^{n-1} f(x))$ giving

$$\text{if } u \in X \text{ then } \mathrm{Dg}(x)(u) = e_{u, v_1, \ldots, v_n}(D^{n-1} f(x)) \tag{14.86}$$

Define now $\mathrm{Dg}(*)(u): U \to Y$ by $x \to (\mathrm{Dg}(*)(u))(x) = \mathrm{Dg}(x)(u)$ then we have by 14.86 that $\mathrm{Dg}(*)(u) = e_{u, v_1, \ldots, v_n} \circ D^{n-1} f$ so that by the chain rule we have $D(\mathrm{Dg}(*)(u))(x) = \mathrm{De}_{u, v_1, \ldots, v_{n-2}}(D^{n-1} f(x)) \circ D(D^{n-1} f)(x) = e_{u, v_1, \ldots, v_n} \circ D^n f(x)$ giving

$$D(\mathrm{Dg}(*)(u))(x) = e_{u, v_1, \ldots, v_{n-2}} \circ D^n f(x) \tag{14.87}$$

Using 14.73 we have given $v \in X$ that $D(\mathrm{Dg}(*)(u))(x)(v) = D^2 g(x)(u)(v)$ so that using 14.87 we have $D^2 g(x)(u)(v) = e_{u, v_1, \ldots, v_n}(D^n f(x)(v)) = (D^n f(x)(v))(u, v_1, \ldots, v_{n-2}) \underset{12.206}{=} D^n f(x)(v: u: v_1, \ldots, v_{n-2})$ □

Lemma 14.117. *Let $n \in \mathbb{N} \vdash n > 1$, $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be normed spaces, $U \subseteq X$ open, $v_1, \ldots, v_{n-1} \in X^{n-1}$, a $f: U \to Y$ that is C^n and $h: U \to Y$ defined by $x \to h(x) = D^{n-1} f(x)(v_1: \ldots v_{n-1})$ then h is C^1 and $\mathrm{Dh}(x)(u) = D^n f(x)(u: v_1: \ldots: v_{n-1})$*

Proof. As $h(x) = D^{n-1}f(x)(v_1\colon \ldots\colon v_{n-1}) = e_{v_1,\ldots,v_{n-1}}(D^{n-1}f(x))$ we have that $h = e_{v_1,\ldots,v_{n-1}} \circ D^{n-1}f$ so using the Chain Rule we have that h is C^1 and $\mathrm{Dh}(x) = \mathrm{De}_{v_1,\ldots v_{n-1}}(D^{n-1}f(x)) \circ D(D^{n-1}f(x)) = e_{v_1,\ldots,v_{n-1}} \circ D^n f(x)$ giving that $\mathrm{Dh}(x)(u) = e_{v_1,\ldots,v_{n-1}}(D^n f(x)(u)) = (D^n f(x)(u))(v_1\colon \ldots\colon v_{n-1}) \underset{12.206}{=} D^n f(x)(u\colon v_1\colon \ldots\colon v_{n-1})$. □

Now we are finally ready to prove the symmetric nature of the differential

Theorem 14.118. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X\rangle$ be a normed space, $\langle Y, \|\|_Y\rangle$ a Banach space, $U \subseteq X$ open and $f\colon U \rightarrow Y$ is C^n then $\forall x \in U$, $\sigma \in P_n$ and $v = (v_1, \ldots, v_n) \in X$ we have $D^n f(x)(v_{\sigma(1)}\colon \ldots\colon v_{\sigma(n)}) = D^n f(x)(v_1\colon \ldots\colon v_n)$*

Proof. We have to consider the following cases for $n \in \mathbb{N}$

1. $(\boldsymbol{n = 1})$ As $P_1 = \{1_{\{1\}}\}$ the proof is trivial.
2. $(\boldsymbol{n = 2})$ Then the theorem follows from 14.112.
3. $(\boldsymbol{n > 2})$ We use induction to prove the theorem so let $S = \{m \in \{2, \ldots,\} | \text{if } m \leqslant n \text{ then } \forall \sigma \in P_m \text{ and } \forall (v_1, \ldots, v_m) \in X^m \text{ we have } D^m f(x)(v_{\sigma(1)}\colon \ldots\colon v_{\sigma(m)}) = D^m f(x)(v_1\colon \ldots\colon v_m)\}$ then we have:
   a. If $m = 2$ then using (14.112) we have that $D^2 f(x)(v_{\sigma(1)}\colon \ldots\colon v_{\sigma(2)}) = D^2 f(x)(v_1, v_2)$ proving that $2 \in S$.
   b. If $m \in S$ then if $m + 1 \leqslant n$ we have that f is C^{m+1}, given $(v_1, \ldots, v_m) \in X^m$, $u, v \in X$ define $g_{v_1,\ldots,v(m+1)_{-2}}\colon U \rightarrow Y$ by $x \rightarrow D^{(m+1)-2}(x)(v_1, \ldots, v_{(m+1)-2})$ we have by 14.116 that $g_{v_1,\ldots,v_{(m+1)-2}}$ is C^2 and $D^{m+1}f(x)(v\colon u\colon v_1, \ldots, v_{(m+1)-2}) = \mathrm{Dg}_{v_1,\ldots,v_{(m+1)-2}}(x)(u)(v) \underset{14.112}{=} \mathrm{Dg}_{v_1,\ldots,v_{(m+1)-2}}(x)(v)(u) = D^{m+1}f(x)(u\colon v\colon v_1, \ldots, v_{(m+1)-2})$ giving

   $$D^{(m+1)}f(x)(v\colon u\colon v_1\colon \ldots\colon v_{(m+1)-2}) = D^{(m+1)}f(x)(u\colon v\colon v_1\colon \ldots\colon v_{(m+1)-2}) \tag{14.88}$$

   Define now $h_{v_1,\ldots v_{(m+1)-1}}\colon U \rightarrow Y$ defined by $x \rightarrow h_{v_1,\ldots,v_{(m+1)-1}}(x) = D^{(m+1)-1}f(x)(v_1\colon \ldots\colon v_{(m+1)-1}) = D^m f(x)(v_1\colon \ldots v_m)$ and given $\sigma \in P_{(m+1)-1} = P_m$ define $h_{v_{\sigma(1)},\ldots v_{\sigma((m+1)-1)}}\colon U \rightarrow Y$ defined by $x \rightarrow h_{v_{\sigma(1)},\ldots,v_{\sigma((m+1)-1)}}(x) = D^{(m+1)-1}f(x)(v_{\sigma(1)}, \ldots, v_{\sigma((m+1)-1)}) = D^m f(x)(v_1\colon \ldots\colon v_m)$ then as $m \in S$ we have $\forall x \in U$ that $h_{v_1,\ldots,v_{(m+1)-1}}(x) = D^m f(x)(v_1\colon \ldots\colon v_m) = D^m f(x)(v_{\sigma(1)}\colon \ldots\colon v_{\sigma(m)}) = h_{v_{\sigma(1)},\ldots,v_{\sigma((m+1)-1)}}(x)$ proving that $h_{v_1,\ldots,v_{(m+1)-1}} = h_{v_{\sigma(1)},\ldots,v_{\sigma((m+1)-1)}}$ then differentiating and using 14.117 we have $D^{m+1}f(x)(u\colon v_1, \ldots, v_m) = \mathrm{Dh}_{v_1,\ldots,v_{(m+1)-1}}(x)(u) = \mathrm{Dh}_{v_{\sigma(1)},\ldots,v_{\sigma((m+1)-1)}}(x) = D^{m+1}f(x)(u\colon v_{\sigma(1)}\colon \ldots\colon v_{\sigma(m)})$ giving

   $$\forall \sigma \in P_m \text{ we have } D^{m+1}f(x)(u\colon v_1\colon \ldots\colon v_m) = D^{m+1}f(x)(u\colon v_{\sigma(1)}\colon \ldots\colon v_{\sigma(m)}) \tag{14.89}$$

   Using then 14.114, 14.88 and 14.89 we have then $\forall \sigma \in P_{n+1}$, $\forall x_1, \ldots, x_{n+1} \in X$ we have $D^{(m+1)}f(x)(v_1\colon \ldots\colon v_{n+1}) = D^{(m+1)}f(x)(v_{\sigma(1)}\colon \ldots\colon x_{\sigma(n+1)})$ proving that $m + 1 \in S$

Using induction we have then that $S=\{2,\ldots\}$ proving the theorem in the case $n>2$ □

As a application of the above theorem let's prove the following theorem

Theorem 14.119. *Let $\langle\mathbb{K},\|\|\rangle$ be the real (complex space) equiped with the canonical norm, $\langle Y,\|\|_Y\rangle$ a Banach space, $U\subseteq\mathbb{K}$ open, $n\in\mathbb{N}$ and $f:U\to Y$ a function then f is C^n if and only if the n-th derivate f^n (see $f^{(n)}$) exists and is continuous. Furthermore if $x\in U,v_1,\ldots,v_n\in\mathbb{K}$ then $D^nf(x)(1:\ldots:1)=f^n(x)$ and $D^nf(x)(v_1:\ldots:v_n)=f^n(x)\cdot\prod_{i\in\{1,\ldots,n\}}v_i$*

Proof. We prove this by induction so let $S=\left\{n\in\mathbb{N}|f\text{ is }C^n\text{ at }x\in U\Leftrightarrow f^n\text{exists and is continuous further }D^nf(x)(1:\ldots:1)=f^{(n)}\text{ and }\mathrm{D}f^n(x)(v_1:\ldots:v_n)=f^n(x)\cdot\prod_{i\in\{1,\ldots,n\}}v_i\right\}$ then we have

1. If $n=1$ then by 14.49 we have that f is C^1 if and only if $f^1=f'$ exists on U and is continuous. Further $f'(x)=\mathrm{Df}(x)(1)=D^1f(x)(1)$ from which it follows that $D^1f(x)(v_1)=v_1\cdot D^1f(x)(1)=f'(x)\cdot v_1=f^1(x)\cdot\prod_{i\in\{1,\ldots,\}}v_i$ proving that $1\in S$
2. If $n\in S$ take then $n+1$ then we have
   a. **(f is C^{n+1})** then given $x\in U$ we have $D^{n+1}f(x)=D(D^nf(x))$ so given $\varepsilon>0$ there exists a $\delta>0$ such that if $|h|<\delta$ and $h\in U_x$ then $\|D^nf(x+h)-D^nf(x)-D^{n+1}f(x)\cdot(h)\|_{L^n(X;Y)}\leqslant\varepsilon\cdot|h|$ so that by 12.208 $\|(D^nf(x+h)-D^nf(x)-D^{n+1}f(x)(h))(1:\ldots:1)\|_Y\leqslant\|D^nf(x+h)-D^nf(x)-D^{n+1}f(x)\cdot(h)\|\cdot\prod_{i\in\{1,\ldots,\}}|1|=\|D^nf(x+h)-D^nf(x)-D^{n+1}f(x)\cdot(h)\|\leqslant\varepsilon\cdot|h|$ giving

$$\|(D^nf(x+h)-D^nf(x)-D^{n+1}f(x)(h))(1:\ldots:1)\|_Y\leqslant\varepsilon\cdot|h| \tag{14.90}$$

As $(D^nf(x+h)-D^nf(x)-D^{n+1}f(x)(h))(1:\ldots:1)\underset{12.207}{=}D^nf(x+h)(1:\ldots:1)-D^nf(x)(1:\ldots:1)-D^{n+1}f(x)(h)(1:\ldots:1)\underset{12.206}{=}D^nf(x+h)(1:\ldots:1)-D^nf(x)(1:\ldots:1)-D^{n+1}f(x)(h:1:\ldots:1)\underset{n\in S}{=}f^n(x+h)-f^n(x)-D^{n+1}f(x)(h:1:\ldots:1)\underset{14.118}{=}f^n(x+h)-f^n(x)-D^{n+1}f(x)(1:\ldots:1:h)=f^n(x+h)-f^n(x)-D^{n+1}f(x)(1:\ldots:1)(h)\underset{\text{linearity of differential}}{=}=f^n(x+h)-f^n(x)-h\cdot D^{n+1}f(x)(1:\ldots:1)(1)=f^n(x+h)-f^n(x)-h\cdot D^{n+1}f(x)(1:\ldots:1)$ giving

$$f^n(x+h)-f^n(x)-h\cdot D^{n+1}f(x)(1:\ldots:1)=(D^nf(x+h)-D^nf(x)-D^{n+1}f(x)(h))(1:\ldots:1) \tag{14.91}$$

So if $0<|h|<\delta$ then $\left\|\frac{f^n(x+h)-f^n(x)}{h}-D^{n+1}f(x)(1:\ldots:1)\right\|_Y=\left\|\frac{f^n(x+h)-f^n(x)-h\cdot D^{n+1}f(x)(1:\ldots:1)}{|h|}\right\|_Y\underset{14.91}{=}\frac{\|(D^nf(x+h)-D^nf(x)-D^{n+1}f(x)(h))(1:\ldots:1)\|_Y}{|h|}\leqslant_{14.90}\frac{\varepsilon\cdot|h|}{|h|}=\varepsilon$ proving that $f^{n+1}(x)=f'(f^n)(x)$ exists and is equal to $D^{n+1}f(x)(1:\ldots:1)$ or

$$f^{n+1}(x)=D^{n+1}f(x)(1:\ldots:1) \tag{14.92}$$

To prove continuity note that $D^{n+1}f$ is continuous so if $\varepsilon > 0$ then there exists a $\delta > 0$ such that if $|x-y| < \delta$ and $y \in U$ then $\|D^{n+1}f(x) - D^{n+1}f(y)\|_{L^{n+1}(X;Y)} < \varepsilon$, as $\|f^{n+1}(x) - f^{n+1}(y)\|_Y = \|D^{n+1}f(x)(1:\ldots: 1) - D^{n+1}f(y)(1:\ldots:1)\|_Y = \|(D^{n+1}f(x) - D^{n+1}f(y))(1:\ldots:1)\|_Y \leqslant \|D^{n+1}f(x) - D^{n+1}f(y)\|_{L^{n+1}(X;Y)} < \varepsilon$ proving contiuity.

b. **(f^{n+1} exists and is continuous)** Then $\forall \varepsilon > 0$ there exists a $\delta > 0$ such that if $0 < |h| < \delta$ and $h \in U_x$ we have that $\left\| \frac{f^n(x+h) - f^n(x)}{h} - f^{n+1}(x) \right\|_Y < \varepsilon$ giving after multiplying both sides by h

$$\|f^n(x+h) - f^n(x) - h \cdot f^{n+1}(x)\|_Y < \varepsilon \cdot |h| \tag{14.93}$$

Now as $n \in S$ we have that $f^n(x+h) = D^n f(x)(1:\ldots:1)$ and $f^n(x) = D^n f(x)(1:\ldots:1)$ we have

$$\|D^n f(x+h)(1:\ldots:1) - D^n f(x)(1:\ldots:1) - h \cdot f^{n+1}(x)\|_Y < \varepsilon \cdot |h| \tag{14.94}$$

Define now the multi linear function $k: L(\mathbb{K}^{n+1}, Y) \rightarrow Y$ defined by $(v_1, \ldots, v_{n+1}) \rightarrow k(v_1, \ldots, v_{n+1}) = f^{n+1}(x) \cdot \prod_{i \in \{1,\ldots,n+1\}} |v_i|$ and take then $g = \mathcal{P}_{n+1}^{-1}(k) \in L^{n+1}(\mathbb{K}; Y)$ so that $g(v_1:\ldots:v_{n+1}) = k(v_1,\ldots,v_{n+1}) = f^{n+1}(x) \cdot \prod_{i \in \{1,\ldots,n+1\}} v_i$. This gives that $g(h)(1:\ldots: 1) \underset{12.206}{=} g(h \cdot 1:\ldots:1) = f^{n+1}(x) \cdot h$ which together with 14.94 gives then

$$\|(D^n f(x+h) - D^n f(x) - g(h))(1:\ldots:1)\|_Y < \varepsilon \cdot |h| \tag{14.95}$$

Using 12.209 and 14.95 we have then that $\|D^n f(x+h) - D^n f(x) - g(h)\| < \varepsilon$ proving that $D^n f$ is differentiable at x with $D(D^n f)(x) = g \Rightarrow D^{n+1} f(x) = g$ where $D^{n+1} f(x)(v_1:\ldots:v_{n+1}) = g(v_1:\ldots:v_{n+1}) = f^{n+1}(x) \cdot \prod_{i \in \{1,\ldots,n+1\}} v_i$ proving that f is c^{n+1} and $D^{n+1} f(x)(v_1:\ldots: v_n) = f^{n+1}(x) \cdot \prod_{i \in \{1,\ldots,n+1\}} v_i$.

So by (a) and (b) we have $n+1 \in S$.

By mathematical induction we have then $S = \mathbb{N}$ proving our theorem. □

14.6 Partial Differentiation

First we must extend the concept of $L^n(X;Y)$ as follows

Theorem 14.120. *Given a finite family $\{\langle X_i, \|\|_i \rangle\}_{i \in \{1,\ldots,n\}}$, $n \in \mathbb{N}$ of normed vector spaces then given $i \in \{1, \ldots, n\}$ and $x = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in \prod_{i \in \{1,\ldots,n\} \setminus \{i\}} X_i = X$ then $(x_1, \ldots, x_{i-1}, *, x_{i+1}, \ldots, x_n): \prod_{j \in \{1,\ldots,n\}} X_j \rightarrow X_i$ (see 14.19) is C^∞.*

Proof. As $(x_1, \ldots, x_{i-1}, *, x_{i+1}, \ldots, x_n)(t) = (x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n) + (0_1, \ldots, 0_{i-1}, t, 0_{i+1}, \ldots, 0_n) = (x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n) + (0_1, \ldots, 0_{i-1}, *, 0_{i+1}, \ldots, 0_n)(t)$ so that $(x_1, \ldots, x_{i-1}, *, x_{i+1}, \ldots, x_n) = C_{X,(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n)} + (0_1, \ldots, 0_{i-1}, *, 0_{i+1}, \ldots, 0_n)$ a sum of a constant function (C^∞) and a linear function (see 14.22) (C^∞) and is thus C^∞. $\square$

We define then partial derivates as follows

Definition 14.121. *Let $n \in \mathbb{N}$, $m \in \mathbb{N}$ $\{\langle X_i, \|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a finite family of normed spaces, $\langle Y, \|\|_Y\rangle$ a normed space, $\left\langle \prod_{i\in\{1,\ldots,n\}} X_i, \|\|\right\rangle$ be the normed space $X = \prod_{i\in\{1,\ldots,n\}} X_i$ with the maximum norm $\|\|$ (see 12.79), $U \subseteq X$ a open set and $f: U \to Y$, $(i_1, \ldots, i_m) \in \{1, \ldots, n\}^m$ then we say that f has a partial differential $D^m_{i_!,\ldots,i_m} f: U \to L^m(X_{i_1}, \ldots, X_{i_m}; Y)$ on U iff*

1. *If $m = 1$ then $D^1_{i_1\ldots,i_i} f$ is a partial differential of f if $\forall x = (x_1, \ldots, x_n) \in U$ there exists a i-partial differential differential $D_{i_1} f(x) \in L(X_i, Y) = L^1(X_i; Y)$ (see 14.30). We define then $D^1_{i_1,\ldots,i_1} f: U \to L^1(X_i; Y)$ by $x \to D^1_{i_1,\ldots,i_1} f(x) = D_{i_1} f(x)$*
2. *If $m > 1$ then f a has a partial differential $D^m_{i_1,\ldots,i_m} f: U \to L^m(X_{i_1}, \ldots, X_{i_m}; Y)$ iff*
   a. *f has a partial differential $D^{m-1}_{i_2,\ldots,i_m} f: U \to L^{m-1}(X_{i_2}, \ldots, X_{i_m}; Y)$ on U*
   b. *$\forall x = (x_1, \ldots, x_n) \in U$ $D^{m-1}_{i_2,\ldots,i_m} f$ has a i-partial differential $D_{i_1}(D^{m-1}_{i_2,\ldots,i_m} f)(x)$. We define then $D^m_{i_1,\ldots,i_m} f: U \to L^m(X_{i_1}, \ldots, X_{i_m}; Y) = L(X_{i_1}, L(X_{i_2}, \ldots, X_{i_m}; Y))$ by $x \to D^m_{i_1,\ldots,i_m} f(x) = D^1_{i_1}(D^{m-1}_{i_2,\ldots,i_m} f)(x)$*

Notation 14.122. *If $m = 1$ then we write $D^1_{i,\ldots,i} f$ as $D^1_i f$*

The next theorem shows that the partial differential on U exists if f is differentiable on U

Theorem 14.123. *Let $n \in \mathbb{N}$, $m \in \mathbb{N}$ $\{\langle X_i, \|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a finite family of normed spaces, $\langle Y, \|\|_Y\rangle$ a normed space, $\left\langle \prod_{i\in\{1,\ldots,n\}} X_i, \|\|\right\rangle$ be the normed space $X = \prod_{i\in\{1,\ldots,n\}} X_i$ with the maximum norm $\|\|$ (see 12.79), $U \subseteq X$ a open set and $f: U \to Y$ a 1 times differentiable function then $\forall i \in \{1, \ldots, n\}$ we have that $D^1_i f: U \to L^1(X_i; Y) = L(X_i, Y)$ exists [and is continuous if f is C^1], further $\forall x \in U$ we have $D^1_i f(x) = D^1 f(x) \circ (0_1, \ldots, 0_{i-1}, *, 0_{i+1}, \ldots, 0_n)$ and $D^1 f(x) = \sum_{i\in\{1,\ldots,n\}} D^1_i f(x) \circ \pi_i$*

Proof. As f is 1-times differentiable on U (or C^1] we have $\forall i \in \{1, \ldots, n\}, x = (x_1, \ldots, x_n) \in U$ by the Chain Rule (see 14.68) and 14.120 that $f \circ (x_1, \ldots, x_{i-1}, *, x_{i+1}, \ldots, x_n)$ is 1-times differentiable (or C^1) on U. By 14.31 we have then that $D_i f(x) = D(f \circ (x_1, \ldots, x_{i-1}, *, x_{i+1}, \ldots, x_n))(x_i) = \mathrm{Df}(x) \circ (0_1, \ldots, 0_{i-1}, *, 0_{i+1}, \ldots, 0_n)$ exists so that $D^1_i f: U \to L(X_i, Y)$ is defined (and exists) by $x \to D^1_i f(x) = D_i f(x) = \mathrm{Df}(x) \circ (0_1, \ldots, 0_{i-1}, *, 0_{i+1}, \ldots, 0_n) = D^1 f(x) \circ (0_1, \ldots, 0_{i-1}, *, 0_{i+1}, \ldots, 0_n)$ and also that $D^1_i f$ is continuous if f is C^1 [as $D^1_i f = D(f \circ (x_1, \ldots, x_{i-1}, *, x_{i+1}, \ldots, x_n))$]. Also by 14.31 we have that $D^1 f(x) = \mathrm{Df}(x) = \sum_{i\in\{1,\ldots,n\}} D_i f(x_i) \circ \pi_i = \sum_{i\in\{1,\ldots,n\}} D^1_i f(x) \circ \pi_i$ $\square$

The next theorem represents a variation on the chain rule

Theorem 14.124. *Let $n\in\mathbb{N}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a finite family of normed spaces, $\langle Y,\|\|_Y\rangle$, $\langle Z,\|\|_Z\rangle$ normed space, $\left\langle \prod_{i\in\{1,\ldots,n\}} X_i, \|\|\right\rangle$ be the normed space $X=\prod_{i\in\{1,\ldots,n\}} X_i$ with the maximum norm $\|\|$ (see 12.79), $U\subseteq\prod_{i\in\{1,\ldots,n\}} X_i$ open and $g\colon U\to Y$ a differentiable function on U, $g(U)\subseteq V$ a open set in Y, $f\colon U\to Z$ a differentiable function on V (so that $f\circ g$ is differentiable and has partial differentials) then $\forall i\in\{1,\ldots,n\}$ and $\forall x=(x_1,\ldots,x_n)\in U$ we have that $D_i(f\circ g)(x_1,\ldots,x_n)=\mathrm{Df}(g(x_1,\ldots,x_n))\circ D_i g(x_1,\ldots,x_n)$*

Proof. Using the definition of a partial differential we have $D_i(f\circ g)(x_1,\ldots,x_n)=D((f\circ g)\circ(x_1,\ldots,x_{i-1},*,x_{i+1},\ldots,x_n))(x_i)=D(f\circ(g\circ(x_1,\ldots,x_{i-1},*,x_{i+1},\ldots,x_n)))(x_i)\underset{\text{Chain rule}}{=}Df((g\circ(x_1,\ldots,x_{i-1},*,x_{i+1},\ldots,x_n))(x_i))\circ D(g\circ(x_1,\ldots,x_{i-1},*,x_{i+1},\ldots,x_n))(x_i)=Df(g(x_1,\ldots,x_n))\circ D_i g(x_1,\ldots,x_n)$ □

The next theorem shows that we can have the reverse of the above theorem if the partial differentials are continuous.

Theorem 14.125. *Let $n\in\mathbb{N}$, $m\in\mathbb{N}$ $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a finite family of normed spaces over $\mathbb{R}$, $\langle Y,\|\|_Y\rangle$ a Banach space over $\mathbb{R}$, $\left\langle \prod_{i\in\{1,\ldots,n\}} X_i, \|\|\right\rangle$ be the normed space $X=\prod_{i\in\{1,\ldots,n\}} X_i$ with the maximum norm $\|\|$ (see 12.79), $U\subseteq X$ a open set and $f\colon U\to Y$ a function so that $\forall i\in D_i^1 f$ exists on U and is continuous then f is C^1.*

Proof. Take $x=(x_1,\ldots,x_n)\in U$ then as $D^1 f_i$ is continous on U we have

$$\forall\varepsilon>0\vDash\exists\delta_{\varepsilon,i}\,\text{with}\,\text{if}\;y\in B_{\|\|}(x,\delta_{\varepsilon,i})\subseteq U\;\text{then}\;\|D_i^1 f(x)-D_i^1 f(y)\|_{L(X_i,Y)}<\varepsilon \tag{14.96}$$

by taking in the above $\delta_\varepsilon=\min(\delta_{\varepsilon,1},\ldots,\delta_{\varepsilon,i})$ we have

$$\forall i\in\{1,\ldots,n\}\text{ and }\forall\varepsilon>0\vDash\exists\delta_\varepsilon\text{ with if }y\in B_{\|\|}(x,\delta_\varepsilon)\subseteq U\text{ then }\|D_i^1 f(x)-D_i^1 f(y)\|_{L(X,Y)}<\varepsilon \tag{14.97}$$

Take now $h=(h_1,\ldots,h_n)\in B_{\|\|}(0,\delta_\varepsilon)$ then we have

$$\|h\|=\max(\{\|h_i\|_i|i\in\{1,\ldots,n\}\}) \tag{14.98}$$

Using 14.23 we have that

$$x+h=x+\sum_{i\in\{1,\ldots,n\}}(0_1,\ldots,0_{i-1},*,0_{i+1},\ldots,0_n)(h_i) \tag{14.99}$$

If we define now $z_0=x$, $z_j=x+\sum_{i\in\{1,\ldots,j\}}(0_i,\ldots,0_{i-1},*,0_{i+1},\ldots,0_n)(h_i)$ if $j>0$ then we have that $\|z_j-x\|=\left\|\sum_{i\in\{1,\ldots,j\}}(0_i,\ldots,0_{i-1},*,0_{i+1},\ldots,0_n)\cdot(h_i)\right\|=\|(h_1,\ldots,h_{j-1},h_j,0_{j+1},\ldots,0_n)\|\leqslant\|h\|$ (see 14.98) so that

$$z_j=\begin{cases} x\text{ if }j=0\\ x+\sum_{i\in\{1,\ldots,j\}}(0_1,\ldots,0_{i-1},*,0_{i+1},\ldots,0_n)(h_i)\text{ if }j\neq 0\end{cases}\in B_{\|\|}(x,\delta_\varepsilon)\subseteq U\text{ if }\|h\|<\delta_\varepsilon \tag{14.100}$$

Next we have that $f(x+h)-f(x)\underset{14.99}{=}f(z_n)-f(z_0)\underset{10.27}{=}\sum_{j\in\{1,\ldots,n\}}(f(z_j)-f(z_{j-1}))$

$$f(x+h)-f(x)\underset{10.27}{=}\sum_{j\in\{1,\ldots,n\}}(f(z_j)-f(z_{j-1}))\tag{14.101}$$

where each of the terms are well defined by 14.100 on $B_{\|\|}(x,\delta_\varepsilon)$. If we define now

$$R(h)=f(x+h)-f(x)-\sum_{i\in\{1,\ldots,n\}}D_i^1 f(x)\circ\pi_i(h)$$

then we have by 14.101 that $R(h)=\sum_{j\in\{1,\ldots,n\}}(f(z_j)-f(z_{j-1}))-\sum_{i\in\{1,\ldots,n\}}D_i^1 f(x)(h_i)=\sum_{j\in\{1,\ldots,n\}}(f(z_j)-f(z_{j-1})-D_j^1 f(x)\cdot h_j)$ giving

$$R(h)=\sum_{j\in\{1,\ldots,n\}}(f(z_j)-f(z_{j-1})-D_j^1 f(x)\cdot h_j)\tag{14.102}$$

Now as $z_1=x+(*,0_2,\ldots,0_n)$ and if $j\neq 0$ we have that $z_j=x+\sum_{i\in\{1,\ldots,,j-1\}}(0_1,\ldots,0_{i-1},*,0_{i+1},\ldots,0_n)(h_i)+(0_1,\ldots,0_{j-1},*,0_{j+1},\ldots,0_n)(h_j)=z_{j-1}+(0_1,\ldots,0_{j-1},*,0_{j+1},\ldots,0_n)(h_j)$ giving togehter with 14.102 that

$$R(h)=\sum_{j\in\{1,\ldots,n\}}(f(z_{j-1}+(0_1,\ldots,0_{j-1},*,0_{j+1},\ldots,0_n)(h_j))-f(z_{j-1})-D_j^1 f(x)(h))\tag{14.103}$$

Define now $\varphi_j\colon B_{\|\|_j}(0,\delta_\varepsilon)\to Y$ by $k\to\varphi_j(k)=f(z_{j-1}+(0_1,\ldots,0_{j-1},*,0_{j+1},\ldots,0_n)(k))-f(z_{j-1})-D_j^1 f(x)(k)$ [which is well defined as $z_{j-1}\in B_{\|\|}(x,\delta_\varepsilon)$ by 14.100, and $\|z_{j-1}+(0_1,\ldots,0_{j-1},*,0_{j+1},\ldots,0_n)(k)-x\|=\left\|\sum_{i\in\{1,\ldots,j-1\}}(0_1,\ldots,0_{i-1},*,0_{i+1},\ldots,0_n)(h_i)+(0_1,\ldots,0_{j-1},*,0_{j+1},\ldots,0_n)(k)\right\|=\|(h_1,\ldots,h_{j-1},k,0_{j+1},\ldots,0_n)\|=\max(\{\|\pi_i((h_1,\ldots,h_{j-1},k,0_{j+1},\ldots,0_n))\|_i|i\in\{1,\ldots,n\}\})<\delta_\varepsilon$ so that $z_{j-1}+(0_1,\ldots,0_{j-1},*,0_{j+1},\ldots,0_n)(k)\in B_{\|\|}(x,\delta_\varepsilon)$]. Note also that $\varphi_j(0)=f(z_{j-1}+(0_1,\ldots,0_{j-1},*,0_{j+1},\ldots,0_n)(0))-f(z_{j-1})-D_j^1 f(x)(0)=0$. We have then that

$$R(h)=\sum_{j\in\{1,\ldots,n\}}\varphi_j(h_j)\,,\varphi_j(0)=0\tag{14.104}$$

Further $\varphi_j(k)=f((z_{j-1})_1,\ldots,(z_{j-1})_{j-1},k,(z_{j-1})_{j+1},\ldots,(z_{j-1})_n)-f(z_{j-1})-D_j^1 f(x)(k)=(f\circ((z_{j-1})_1,\ldots,(z_{j-1})_{j-1},*,(z_{j-1})_{j+1},\ldots,(z_{j-1})_n))(k)-f(z_{j-1})-D_j^1 f(x)(k)$ which is differentiable as $(f\circ((z_{j-1})_1,\ldots,(z_{j-1})_{j-1},*,(z_{j-1})_{j+1},\ldots,(z_{j-1})_n))$ is C^1 (see 14.30 and the fact that by the assumption we have that there is a j-partial differential on U exists and is continuous on U), that a constant is C^∞, that the continuous linear function $D_j^1 f(x)$ is C^∞ and that the sum of C^1 functions is C^1. Further $D\varphi_j(k)=D(f\circ((z_{j-1})_1,\ldots,(z_{j-1})_{j-1},*,(z_{j-1})_{j+1},\ldots,(z_{j-1})_n))(k)-D_j^1 f(x)\underset{14.30}{=}D_j f((z_{j-1})_1,\ldots,(z_{j-1})_{j-1},k,(z_{j-1})_{j-1},\ldots,(x_{j-1})_n)-D_j f(x)$. Using 14.97 we have then that $\|D\varphi_j(k)\|_{L(X_j,Y)}=\|D_j f((z_{j-1})_1,\ldots,(z_{j-1})_{j-1},k,(z_{j-1})_{j-1},\ldots,(x_{j-1})_n)-D_j f(x)\|_{L(X_j,Y)}<\varepsilon$ proving thus

$$\forall{=}j\in\{1,\ldots,n\}\text{ we have }\varphi_j\text{ is }C^1\text{ and }\forall k\in B_{\|\|_j}(0,\delta_\varepsilon)\text{ we have }\|D\varphi_j(k)\|_{L(X_j,Y)}<\varepsilon$$

Using then the Mean Value Theorem (see 14.102) we have as $h_j, 0 \in B_{\|\|_j}(0, \delta_\varepsilon)$ a convex set that $\|\varphi_j(h_j)\|_Y = \|\varphi_j(h_j) - \varphi_j(0)\|_Y \leqslant \varepsilon \cdot \|h_j - 0\|_j = \varepsilon \cdot \|h_j\|_j$, so we have that $\|R(h)\|_Y = \left\|\sum_{j \in \{1, \ldots, n\}} \varphi_j(h_j)\right\|_Y \leqslant \sum_{j \in \{1, \ldots, n\}} \|\varphi_j(h_j)\| < \sum_{j \in \{1, \ldots, n\}} \varepsilon \cdot \|h_j\|_j = \varepsilon \cdot \sum_{j \in \{1, \ldots, n\}} \|h\|_j \leqslant n \cdot \varepsilon \cdot \|h\|$ proving that

$$\text{if } h \in B_{\|\|}(0, \delta_\varepsilon) \text{ then } \|R(h)\|_Y < n \cdot \varepsilon \cdot \|h\|$$

Take now $\xi > 0$ then if we have that $\|h\| < \delta_{\frac{\xi}{n}}$ then $\left\|f(x+h) - f(x) - \sum_{i \in \{1, \ldots, n\}} D_i^1 f(x) \circ \pi_i(h)\right\|_Y = \|R(h)\|_Y \leqslant n \cdot \frac{\xi}{n} \cdot \|h\| = \xi \cdot \|h\|$ proving that f is differentiable on $x \in U$ with differential $\mathrm{Df}(x) = \sum_{i \in \{1, \ldots, n\}} D_i^1 f(x) \circ \pi_i$. Now if $\xi > 0$ then if $y \in B_{\|\|}\left(x, \delta_{\frac{\xi}{N}}\right)$ $\|\mathrm{Df}(x) - \mathrm{Df}(y)\|_{L(X,Y)} = \left\|\sum_{i \in \{1, \ldots, n\}} D_i^1 f(x) \circ \pi_i - \sum_{i \in \{1, \ldots, n\}} D_i^1 f(y) \circ \pi_i\right\|_{L(X,Y)} = \left\|\sum_{i \in \{1, \ldots, n\}} (D_i^1 f(x) - D_i^1 f(y)) \circ \pi_i\right\|_{L(X,Y)} \leqslant \sum_{i \in \{1, \ldots, n\}} \|(D_i^1 f(x) - D_i^1 f(y)) \circ \pi_i\|_{L(X,Y)} \leqslant \sum_{i \in \{1, \ldots, n\}} \|(D_i^1 f(x) - D_i^1 f(y))\|_{L(X_i,Y)} \cdot \|\pi_i\|_{L(X,X_i)} \leqslant_{12.182} \sum_{i \in \{1, \ldots, n\}} \|D_i^1 f(x) - D_i^1 f(y)\|_{L(X_i,Y)} \leqslant_{14.97} n \cdot \frac{\xi}{n} = \xi$ proving continuity of $\mathrm{Df}(x)$ which means that f is C^1 □

Theorem 14.126. *Let $n \in \mathbb{N}$ $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1, \ldots, n\}}$ a finite family of normed spaces, $\langle Y, \|\|_Y\rangle$ a normed space, $\left\langle \prod_{i \in \{1, \ldots, n\}} X_i, \|\|\right\rangle$ be the normed space $X = \prod_{i \in \{1, \ldots, n\}} X_i$ with the maximum norm $\|\|$ (see 12.79), $U \subseteq X$ a open set and $f: U \rightarrow Y$ a r-times differentiable function [C^r function] then $\forall i = (i_1, \ldots, i_r) \in \{1, \ldots, n\}^r$ we have that $D^r_{i_1, \ldots, i_r} f: U \rightarrow L^r(X_{i_1}, \ldots, X_{i_r}; Y)$ exists on U [and is continuous function if f is C^r], further $\forall x \in U$ and $\{h_j\}_{j \in \{1, \ldots, r\}}$ with $h_j \in X_{i_j}$ we have $D^r_{i_1, \ldots, i_r} f(x)(h_1: \ldots: h_r) = D^r f(x)((0_1, \ldots, 0_{i_1 - 1}, h_1, 0_{i_1+1}, \ldots, 0_n): \ldots: (0_1, \ldots, 0_{i_{1_r} - 1}, h_1, 0_{i_r+1}, \ldots, 0_n))$*

Proof. We prove this by induction so let $S = \{s \in \mathbb{N} | \text{if } s \leqslant r \text{ then } \forall (i_1, \ldots, i_s) \in \{1, \ldots, n\}^s$ we have that $D^s_{i_1, \ldots, i_s} f$ exists on U [and is continuous if f is C^s] further $\forall x \in U$ and $\{h_j\}_{j \in \{1, \ldots, s\}}$ with $h_j \in X_{i_j}$ we have $D^s_{i_1, \ldots, i_s} f(x)(h_1: \ldots: h_s) = D^s f(x)((0_1, \ldots, 0_{i_1-1}, h_1, 0_{i_1+1}, \ldots, 0_n): \ldots: (0_1, \ldots, 0_{i_s - 1}, h_s, 0_{i_s+1}, \ldots, 0_n))\}$ then we have:

1. If $s = 1$ use then 14.123 to prove that $\forall (i_1) \in \{1, \ldots, n\}^1 \Rightarrow i_1 \in \{1, \ldots, n\}$ we have that $D^1_{i_1} f$ exists on U , $\forall x \in U$ we have $D^1_{i_1} f(x) = D^1 f(x) \circ (0_1, \ldots, 0_{i_1-1}, *, 0_{i_1+1}, \ldots, 0_n) \Rightarrow \forall \{h_i\}_{i \in \{1, \ldots, 1\}}$ with $h_1 \in X_{i_1}$ we have $D^1_{i_1} f(x)(h_1: \ldots h_1) = D^1_{i_1} f(x)(h_1) = D^1 f(x)(0_1, \ldots, 0_{i_1-1}, h_1, 0_{i_1+1}, \ldots, 0_n)$ proving that $1 \in S$.

2. If $s \in S$ take then $s+1$ and assume that $s+1 \leqslant r$ then if $(i_1, \ldots, i_s) \in \{1, \ldots, n\}^s$ we define $\psi_{i_1, \ldots, i_s}: \prod_{j \in \{1, \ldots, s\}} X_{i_j} \rightarrow X^s$ by $(h_1, \ldots, h_s) \rightarrow \psi_{i_1, \ldots, i_s}(h_1, \ldots, h_s) = ((0_1, \ldots, 0_{i_1-1}, h_1, 0_{i_1+1}, \ldots, 0_n), \ldots, (0_!, \ldots, 0_{i_s-1}, h_s, 0_{i_s+1}, \ldots, 0_n))$. Then if $\alpha, \beta \in \mathbb{K}$ and $(h_1, \ldots, h_s), (g_1, \ldots, g_s)$ we have that $\psi_{i_1, \ldots, i_s}(\alpha \cdot (h_1, \ldots, h_s) + \beta \cdot (h_1, \ldots, h_s)) = \psi_{i_1, \ldots, i_s}(\alpha \cdot h_1 + \beta \cdot g_1, \ldots, \alpha \cdot h_s + \beta \cdot g_s) = ((0_1, \ldots, 0_{i_1-1}, \alpha \cdot h_1 + \beta \cdot g_1, 0_{i_1+1}, \ldots, 0_n) \ldots (0_1, \ldots, 0_{i_s-1}, \alpha \cdot h_s + \beta \cdot g_s, 0_{i_s+1}, \ldots, 0_n)) = ((\alpha \cdot 0_1 + \beta \cdot 0_1, \ldots, \alpha \cdot 0_{i_1-1} + \beta \cdot 0_{i_1+1}, \alpha \cdot h_1 + \beta \cdot g_1, \alpha \cdot 0_{i_1+1} + \beta \cdot 0_{i_1+1}, \ldots, 0_n) \ldots (\alpha \cdot 0_1 + \beta \cdot 0_1, \ldots, \alpha \cdot 0_{i_s-1} + \beta \cdot 0_{i_s-1}, \alpha \cdot h_s + \beta \cdot g_s, \alpha \cdot 0_{i_s+1} + \beta \cdot 0_{i_s+1}, \ldots, \alpha \cdot 0_n + \beta \cdot 0_n)) = (\alpha \cdot (0_1, \ldots, 0_{i_1-1}, h_1, 0_{i_1+1}, \ldots, 0_n) + \beta \cdot (0_1, \ldots, 0_{i_1-1}, g_1, 0_{i_1+1}, \ldots, 0_n), \ldots, \alpha \cdot$

$(0_1 \ldots, 0_{i_s-1}, h_s, 0_{i_s+1}, \ldots, 0_n) + \beta \cdot (0_1, \ldots, 0_{i_s-1}, g_s, 0_{i_s+1}, \ldots, 0_n)) = \alpha \cdot ((0_1, \ldots, 0_{i_1-1}, h_i, 0_{i_1+1}, \ldots, 0_n), \ldots, (0_1, \ldots, 0_{i_s-1}, h_s, 0_{i_s+1}, \ldots, 0_n)) + \beta \cdot ((0_1, \ldots, 0_{i_1-1}, g_i, 0_{i_1+1}, \ldots, 0_n), \ldots, (0_1, \ldots, 0_{i_s-1}, g_s, 0_{i_s+1}, \ldots, 0_n)) = \alpha \cdot \psi_{i_1,\ldots,i_s}(h_1, \ldots, h_s) + \beta \cdot \psi_{i_1,\ldots,i_s}(h_1, \ldots, h_s)$ proving that:

$$\psi_{i_1,\ldots,i_s} \text{ is linear} \tag{14.105}$$

Also we have that $\|\psi_{i_1,\ldots,i_s}(h_1, \ldots, h_s)\|_{X^s} = \|((0_1, \ldots, 0_{i_1-1}, h_1, 0_{i_1+1}, \ldots, 0_n), \ldots, (0_!, \ldots, 0_{i_s-1}, h_s, 0_{i_s+1}, \ldots, 0_n))\|_{X^s} = \max(\{\|(0_1, \ldots, 0_{i_j-1}, h_j, 0_{i_j+1}, \ldots, 0_n)\|_X | j \in \{1, \ldots, s\}\}) \underset{\|\|_X \text{ is maximum norm}}{=} \max(\{\|h_j\|_j | j \in \{1, \ldots, s\}\}) = \|(h_1, \ldots, h_s)\|_{X^s}$ proving that:

$$\psi_{i_1,\ldots,i_s} \text{ is continuous and } \|\psi_{i_1,\ldots,i_s}\|_{L(\prod_{j \in \{1,\ldots,s\}}, X^s)} \leqslant 1 \tag{14.106}$$

Define now $\varphi_{i_1,\ldots,i_s}: L^s(X; Y) \to L^s(X_{i_1}, \ldots, X_{i_s}; Y)$ by $L \to \varphi_{i_1,\ldots,i_s}(L) = \mathcal{P}_s(L) \circ \psi_{i_!,\ldots,i_s}$ then we have that $\forall (h_1, \ldots, h_s) \in L^s(X_{i_1}, \ldots, X_{i_s}; Y)$ we have that $\varphi_{i_1,\ldots,i_s}(L)(h_1, \ldots, h_s) = \mathcal{P}_s(L)(\psi_{i_1,\ldots,i_s}(h_1, \ldots, h_s)) = \mathcal{P}_s(L)((0_1, \ldots, 0_{i_1-1}, h_1, 0_{i_1+1}, \ldots, 0_n), \ldots, (0_!, \ldots, 0_{i_s-1}, h_s, 0_{i_s+1}, \ldots, 0_n)) = L((0_1, \ldots, 0_{i_1-1}, h_1, 0_{i_1+1}, \ldots, 0_n) : \ldots : (0_!, \ldots, 0_{i_s-1}, h_s, 0_{i_s+1}, \ldots, 0_n))$ proving that

$$\forall (h_1, \ldots, h_s) \in L^s(X_{i_1}, \ldots, X_{i_s}; Y) \vDash \varphi_{i_1,\ldots,i_s}(L)(h_1, \ldots, h_s) = L((0_1, \ldots, 0_{i_1-1}, h_1, 0_{i_1+1}, \ldots, 0_n) : \ldots : (0_!, \ldots, 0_{i_s-1}, h_s, 0_{i_s+1}, \ldots, 0_n)) \tag{14.107}$$

Now if $\alpha, \beta \in \mathbb{K}$ and $L_1, L_2 \in L^s(X; Y)$ then we have that $\varphi_{i_1,\ldots,i_s}(\alpha \cdot L_1 + \beta \cdot L_2) = \mathcal{P}_s(\alpha \cdot L_1 + \beta \cdot L_2) \circ \psi_{i_1,\ldots,i_s} \underset{12.213}{=} (\alpha \cdot \mathcal{P}_s(L_1) + \beta \cdot \mathcal{P}_s(L_2)) \circ \psi_{i_1,\ldots,i_s} = \alpha \cdot \mathcal{P}_s(L_1) \circ \psi_{i_1,\ldots,i_s} + \beta \cdot \mathcal{P}_s(L_2) \circ \psi_{i_1,\ldots,i_s} = \alpha \cdot \varphi_{i_1,\ldots,i_s}(L_1) + \beta \cdot \varphi_{i_1,\ldots,i_s}(L_2)$ proving that:

$$\varphi_{i_1,\ldots,i_s} \text{ is linear} \tag{14.108}$$

Next we have that $\|\varphi_{i_1,\ldots,i_s}(L)\|_{L^s(X_{i_1},\ldots,X_{i_s};Y)} = \|\mathcal{P}_S(L) \circ \psi_{i_1,\ldots,i_s}\|_{L^s(X_{i_1},\ldots,X_{i_s};Y)} \leqslant_{12.183} \|\mathcal{P}_S(L)\|_{L(X^s;Y)} \cdot \|\psi_{i_1,\ldots i_s}\|_{L(\prod_{j \in \{1,\ldots,s\}}, X^s)} \leqslant_{14.106} \|\mathcal{P}_S(L)\|_{L(X^s;Y)} \leqslant_{12.213} \|L\|_{L^s(X;Y)}$ proving that

$$\varphi_{i_1,\ldots,i_s} \text{ is continuous} \tag{14.109}$$

Now we have $\varphi_{i_1,\ldots,i_s}(D^s f(x))(h_1 : \ldots : h_s) \underset{14.107}{=} D^s f(x)((0_1, \ldots, 0_{i_1-1}, h_1, 0_{i_1+1}, \ldots, 0_n) : \ldots : (0_1, \ldots, 0_{i_s-1}, h_s, 0_{i_s+1}, \ldots, 0_n)) \underset{s \in S}{=} D^s_{i_!,\ldots,i_s} f(x)(h_1 : \ldots : h_s)$ so that by 12.205 we have $\varphi_{i_1,\ldots,i_s}(D^s f(x)) = D^s_{i_1,\ldots,i_s} f(x)$ or:

$$\varphi_{i_1,\ldots,i_s} \circ D^s f = D^s_{i_1,\ldots,i_s} f \tag{14.110}$$

Now as $\varphi_{i_1,\ldots,i_1}$ is linear and continuous (see 14.108 and 14.109) and thus C^∞ and as $s + 1 \leqslant r$ we have that $D^s f$ is differentiable [and continuous if f is C^r] we have that $D^s_{i_1,\ldots,i_s} f$ is differentiable [or C^1 if f is C^r] and if $x \in U$ then $D(D^s_{i_1,\ldots,i_s} f)(x) = D\varphi_{i_1,\ldots,i_n}(D^s f(x)) \circ D(D^s f)(x) = \varphi_{i_1,\ldots,i_s} \circ D^{s+1} f(x)$ so if $i_0 \in \{1, \ldots, n\}$ we have that $D_{i_0}(D^s_{i_1,\ldots,i_S} f)(x) \underset{14.31}{=} D(D^s_{i_1,\ldots,i_s} f)(x) \circ$

$(0_1, \ldots, 0_{i_{s+1}-1}, *, 0_{i_{s+1}+1}, \ldots, 0_n) = \varphi_{i_1,\ldots,i_s} \circ D^{s+1} f(x) \circ (0_1, \ldots, 0_{i_0-1}, *, 0_{i_0+1}, \ldots, 0_n)$ so if $(h_0, \ldots, h_s) \in \prod_{j \in \{0,\ldots,s\}} X_{i_j}$ we have that $D^{s+1}_{i_0,\ldots,i_s} f(x) = (D_{i_0}(D^s_{i_1,\ldots,i_s} f)(x)(h_0))(h_1: \ldots: h_s) = \varphi_{i_1,\ldots,i_s}(D^{s+1} f(x)(0_1, \ldots, 0_{i_0-1}, h_0, 0_{i_0+1}, \ldots, 0_n))(h_1: \ldots: h_1) = \varphi_{i_1,\ldots,i_s}(D^{s+1} f(x)(0_1, \ldots, 0_{i_0-1}, h_0, 0_{i_0+1}, \ldots, 0_n))(h_1: \ldots: h_1) = D^{s+1} f(x)(0_1, \ldots, 0_{i_0-1}, h_0, 0_{i_0+1}, \ldots, 0_n)((0_1, \ldots, 0_{i_1-1}, h_1, 0_{i_1+1}, \ldots, 0_n): \ldots: (0_1, \ldots, 0_{i_s-1}, h_s, 0_{i_s+1}, \ldots, 0_n)) = D^{s+1} f(x)((0_1, \ldots, 0_{i_0-1}, h_0, 0_{i_0+1}, \ldots, 0_n): \ldots: (0_1, \ldots, 0_{i_s-1}, h_s, 0_{i_s+1}, \ldots, 0_n))$ or if we use $\forall i \in \{1, \ldots, n+1\}$ that $j_i = i_{i-1}$ we have that

$$\forall (h_1, \ldots, h_{s+1}) \in \prod_{i \in \{1,\ldots,s+1\}} X_{j_i} \text{ we have } D^{s+1}_{j_1,\ldots,j_{s+1}} f(x) = D^{s+1} f(x)((0_1, \ldots, 0_{j_1-1}, h_1, 0_{j_1+1}, \ldots, 0_n): \ldots: (0_1, \ldots, 0_{j_{s+1}-1}, h_s, 0_{j_{s+1}+1}, \ldots, 0_n)) \quad (14.111)$$

The above proves that $s+1 \in S$

Mathematical induction proves then that $S = \mathbb{N}$ and our theorem □

Let's now define paritial derivates of higher order.

Definition 14.127. *Let $n, m \in \mathbb{N}$, $\langle \mathbb{K}^n, \|\|\rangle$ be the vector space $\mathbb{K}^n$ with any norm, $\langle Y, \|\|_Y\rangle$ a normed space, $U \subseteq K^n$ a open set, $f: U \to Y$ then given $(i_1, \ldots, i_m) \in \{1, \ldots, n\}^m$ we say that f has a* ***partial derivative*** *$\partial^m_{i_1,\ldots,i_m} f$ on U if*

1. *If $m = 1$ then if $(i) \in \{1, \ldots, m\}^1$ we have $\partial^1_i f$ exists on U if $\forall x \in U$ we have that f has a partial derivative at x (see 14.32). We define then $\partial^1_i f: U \to Y$ by $x \to \partial^1_i f(x)$.*
2. *If $m > 1$ then if $(i_1, \ldots, i_m) \in \{1, \ldots, n\}^m$ we say that f has a partial derivative $\partial^m_{i_1,\ldots,i_m} f$ on U if $\partial^{m-1}_{i_2,\ldots,i_m} f$ exists on U and that $\forall x \in U$ the partial derivative $\partial_{i_1}(\partial^{m-1}_{i_2,\ldots,i_n} f)(x)$ exists. We define then $\partial^m_{i_1,\ldots,i_m} f: U \to Y$ by $x \to \partial^m_{i_1,\ldots,i_m} f(x) = \partial_{i_1}(\partial^{m-1}_{i_2,\ldots,i_n} f)(x)$*

Theorem 14.128. *Let $n, m \in \mathbb{N}$, $\langle \mathbb{K}^n, \|\|\rangle$ be the vector space $\mathbb{K}^n$ with any norm, $\langle Y, \|\|_Y\rangle$ a normed space, $U \subseteq K^n$ a open set, $f: U \to Y$ so that given f is r-times differentiable [or C^m] then $\forall (i_1, \ldots, i_r) \in \{1, \ldots, n\}^r$ we have that $\partial^r_{i_1,\ldots,i_r} f$ exists on U [and is continuous on U if f is C^r]. Furthermore $\forall x \in U$ we have $\partial^r_{i_1,\ldots,i_r} f(x) = (D^r_{i_1,\ldots,i_r} f(x))\left(\underbrace{1: \ldots: 1}_{r}\right)$ or $\partial^r_{i_1,\ldots,i_r} f(x) = D^r f(x)(e_{i_1}: \ldots: e_{i_r})$ where $\{e_i\}_{i \in \{1,\ldots,n\}}$ is the canonical base on $\mathbb{K}^n$ (defined by $(e_i)_j = \delta_{i,j}$)*

Proof. We proof this by induction so let $S = \{s \in \mathbb{N} | \text{if } s \leqslant r \text{ then } \forall (i_1, \ldots, i_s)\ \partial^r_{i_1,\ldots,i_r} f \text{ exists on } U \text{ (is continuous if } f \text{ is } C^r), \partial^r_{i_1,\ldots,i_r} f(x) = (D^r_{i_1,\ldots,i_r} f(x))(1: \ldots: 1) = D^r f(x)(e_{i_1}: \ldots: e_{i_r})\}$ then we have:

1. If $s = 1$ then by 14.34 we have that $\forall x \in U$ the partial derivative exists proving that $\partial_i f$ exists on U and $\forall x \in U$ we have $\partial_i f(x) = D_i f(x)(1) = \mathrm{Df}(x)(e_i)$. Using 14.73 we have also that $\mathrm{Df}(*)(e_i) = \partial_i f$ is continuous. So we have $1 \in S$

2. If $s \in S$ then if $s+1 \leqslant r$ we have $s \leqslant r$ and thus $\forall (i_1, \ldots, i_s) \in (1, \ldots, n)^s$ we have that $\forall x \in U\ \partial^s_{i_1, \ldots, i_n} f(x)$ exists and $\partial^s_{i_1, \ldots, i_n} f(x) = D^s f(x)(e_{i_1}: \ldots : e_{i_s})$. Define now $\varphi_{i_!, \ldots, i_s}: L^s(\mathbb{K}^n; Y) \rightarrow Y$ by $L \rightarrow \varphi(L) = L(e_{i_1}: \ldots : e_{i_s})$ then we have by 12.207 that $\varphi_{i_1, \ldots, i_s}$ is linear. As also by 12.208 we have that $\|\varphi_{i_1, \ldots, . i_s}(L)\|_Y = \|L(e_{i_1}: \ldots : e_{i_s})\|_Y \leqslant \|L\| \cdot \prod_{i \in \{1, \ldots, n\}} \|e_{i_1}\|_{\mathbb{K}^n} = \|L\|$ we have that $\varphi_{i_1, \ldots, i_s}$ is continuous and thus C^∞. So we have that $\partial^s_{i_1, \ldots, i_n} f(x) = D^s f(x)(e_{i_1}: \ldots : e_{i_s}) = \varphi_{i_1, \ldots, i_s}(D^s f(x)) \Rightarrow \partial^s_{i_1, \ldots, i_s} f = \varphi_{i_1, \ldots, i_s} \circ D^s f(x)$ which is differentiable (as $s + 1 \leqslant r$) and C^1 if f is C^r so that the differential is continuous if f is C^r] and we have that $\forall x \in U$ that $D(\partial^s_{i_1, \ldots, i_s} f)(x) = D\varphi_{i_1, \ldots, i_s}(D^s f(x)) \circ D(D^s f)(x) = \varphi_{i_1, \ldots, i_s} \circ D^{s+1} f(x)$. Take now $i_0 \in \{1, \ldots, n\}$ then by 14.34 $\partial_{i_0}(\partial^s_{i_1, \ldots, i_s} f)(x)$ exists and $\partial_{i_0}(\partial^s_{i_1, \ldots, i_s} f)(x) \underset{14.34}{=} D(\partial^s_{i_1, \ldots, i_S} f)(e_{i_0}) = \varphi_{i_1, \ldots, i_s}(D^{s+1} f(x)(e_{i_0})) = D^{s+1} f(x)(e_{i_0})(e_{i_!}: \ldots : e_{i_1}) = D^{s+1} f(x)(e_{i_1}: \ldots : e_{i_1})$ proving that

$$\forall x \in U \text{ we have } \partial^{s+1}_{i_0, \ldots i_s} f(x) = D^{s+1} f(x)(e_{i_1}: \ldots : e_{i_1}) \tag{14.112}$$

Define now $\psi: L^s\Big(\underbrace{\mathbb{K}, \ldots, \mathbb{K}}_{s}; Y\Big) \rightarrow Y$ by $L \rightarrow \psi(L) = L\Big(\underbrace{1: \ldots : 1}_{s}\Big)$ then we have by 12.207 and 12.208 again that ψ is linear and continuous. As $\partial^s_{i_1, \ldots, i_s} f(x) \underset{s \in S}{=} (D^s_{i_1, \ldots, i_s} f(x))\Big(\underbrace{1: \ldots : 1}_{r}\Big) = \psi(D^s_{i_1, \ldots, i_s} f(x)) = (\psi \circ D^s_{i_1, \ldots, i_s} f)(x) \Rightarrow \partial^s_{i_1, \ldots, i_s} f = \psi \circ D^s_{i_1, \ldots, i_s} f$ so that $D(\partial^s_{i_1, \ldots, i_s} f)(x) = D\psi(D^s_{i_1, \ldots, i_s} f(x)) \circ D(D^s_{i_1, \ldots, i_s} f)(x) = \psi \circ D(D^s_{i_1, \ldots, i_s} f)(x)$. Now $\partial_{i_0}(\partial^s_{i_1, \ldots, i_s} f)(x) = D_{i_0}(\partial^s_{i_1, \ldots, i_s} f)(x)(1) \underset{14.31}{=} D(\partial^s_{i_1, \ldots, i_s} f)(x)(0_1, \ldots, 0_{i_0 - 1}, 1, 0_{i_0+1}, \ldots 0_n) = (\psi \circ D(D^s_{i_1, \ldots, i_s} f)(x))(0_1, \ldots, 0_{i_0-1}, 1, 0_{i_0+1}, \ldots 0_n) = \psi(D(D^s_{i_1, \ldots, i_s} f)(x)(0_1, \ldots, 0_{i_0-1}, 1, 0_{i_0+1}, \ldots 0_n)) = \psi(D(D^s_{i_1, \ldots, i_s} f)(x)((0_1, \ldots, 0_{i_0-1}, *, 0_{i_0+1}, \ldots, 0_n)(1))) = \psi(D_{i_0}(D^s_{i_1, \ldots, i_n} f)(x)(1)) = \psi(D^{s+1}_{i_0, \ldots, i_s} f(x)(1)) = D^{s+1}_{i_0, \ldots, i_s} f(x)(1)\Big(\underbrace{1: \ldots : 1}_{s}\Big) = D^{s+1}_{i_0, \ldots, i_s} f(x)\Big(\underbrace{1: \ldots : 1}_{s+1}\Big)$ proving

$$\partial^{s+1}_{i_0, \ldots, i_s} f(x) = D^{s+1}_{i_0, \ldots, i_s} f(x)\Big(\underbrace{1: \ldots : 1}_{s+1}\Big) \tag{14.113}$$

Substituting $j_1, \ldots, j_{s+1}$ for $i_0, \ldots, i_s$ in 14.112,14.113 gives then finaly that $s + 1 \in S$.

Mathematical induction proves then our theorem. □

14.7 Inverse function theory

Theorem 14.129. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be normed spaces then a linear continuous isomorphism $L: X \rightarrow Y$ is a toplinear isormorphism (see 12.234) if also L^{-1} is continuous.*

Proof. This is trivial using the definition (see 12.234). □

Theorem 14.130. *If $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ are Banach spaces then if $L: X \rightarrow Y$ is a linear continuous isomorphism (or $L \in L(X, Y)$ and is a isomorphism) then L is a toplinear isomorphism.*

Proof. This is a consequence of 12.409. □

Definition 14.131. *If $\langle X, \|\|\rangle$ is a normed space then $\mathcal{G}L(X) = \{L \in L(X, Y)|L \text{ is toplinear}\} \subseteq L(X, Y)$*

Theorem 14.132. *If $\langle X, \|\|\rangle$ is a Banach space then $\mathcal{G}L(X)$ is a group with the composition operator as the group operator.*

Proof. If $L_1, L_2 \in \mathcal{G}L(X)$ then $L_1 \circ L_2$ is also a isomorphism that is linear and continuous and by 12.409 is continuous as X is a Banach space so $L_1 \circ L_2 \in \mathcal{G}L(X)$ proving that $\circ: X \times X \rightarrow X$ defined by $(L_1, L_2) \rightarrow L_1 \circ L_2$ is well defined.

1. **(identity)** Consider $1_X: X \rightarrow X$ which is evidently linear, continuous and a isomorphism with a continuous inverse. If $L \in \mathcal{G}L(X)$ then $L \circ 1_X = L = 1_X \circ L$
2. **(inverse element)** If $L \in \mathcal{G}L(X)$ then by definition $L^{-1} \in \mathcal{G}L(X)$ and $L \circ L^{-1} = 1_X = L^{-1} \circ L$
3. **(associativity)** This is trivial as composition of functions is associative. □

Definition 14.133. *Let $\langle X, \|\|\rangle$ be a normed space, $H: X \rightarrow X$ a continuous linear function and $n \in \mathbb{N}$ then we define $H^n: X \rightarrow X$ recursively as follows*

$$\begin{aligned} H^1 &= H \\ H^{n+1} &= H \circ H^n \end{aligned}$$

Note 14.134. Let $\langle X, \|\|\rangle$ be a normed space, $H \in L(X, X)$ then $-H \in L(X, X)$ obviously and $\forall n \in \mathbb{N}$ we have $(-H)^n = (-1)^n \cdot H^n$

Proof. Let $S = \{n \in \mathbb{N}|\text{if } H \in L(X, X) \text{ then } (-H)^n = (-1)^n \cdot H\}$ then:

1. If $n = 1$ then $(-H)^1 = -H = (-1)^1 \cdot H$
2. If $n \in S$ then $(-H)^{n+1} = (-H) \circ (-H)^n \underset{n \in S}{=} (-H) \circ (-1)^n \cdot H^n = (-1) \cdot (-1)^n \cdot H \circ H^n = (-1)^{n+1} \cdot H^{n+1}$ proving that $n + 1 \in S$ □

Lemma 14.135. *Let $\langle X, \|\|_X\rangle$ be a normed vector space and $H \in L(X, X)$, $n \in \mathbb{N}$ then $H^n: X \rightarrow X$ is linear and continuous with $\forall x \in X \vDash \|H^n(x)\|_X \leqslant \|H\|^n \|x\|_X$ (or $\|H^n\| \leqslant \|H\|^n$). In addition we have that $\sum_{i=1}^n H^i$ is continous with $\forall x \in X \vDash \|\sum_{i=1}^n H^i(x)\|_X \leqslant (\sum_{i=1}^n \|H\|^i) \cdot \|x\|_X$ (or $\|\sum_{i=1}^n H^i\| \leqslant \sum_{i=1}^n \|H\|^i$)*

Proof.

1. **(linearity)** We use induction so let $S = \{n \in \mathbb{N}|H^n \text{ is linear}\}$ then we have
   a. If $n = 1$ then $H^1(\alpha \cdot x + \beta \cdot y) = H(\alpha \cdot x + \beta \cdot y) \underset{H \text{ is linear}}{=} \alpha \cdot H(x) + \beta \cdot H(y) = \alpha \cdot H^1(x) + \beta \cdot H^1(y) \Rightarrow 1 \in S$
   b. If $n \in S$ then $H^{n+1}(\alpha \cdot x + \beta \cdot y) = H(H^n(\alpha \cdot x + \beta \cdot y)) \underset{n \in S}{=} H(\alpha \cdot H^n(x) + \beta \cdot H^n(y)) \underset{H \text{ is linear}}{=} \alpha \cdot H(H^n(x)) + \beta \cdot H(H^n(y)) = \alpha \cdot H^{n+1}(x) + \beta \cdot H^{n+1}(y)$ proving that $n + 1 \in S$

2. **(continuity of H^n)** Again we use induction so let $S=\{n\in\mathbb{N}|H^n\text{ is continuous and }\forall x\in X\text{ we have }\|H^n(x)\|_X\leqslant\|H\|^n\cdot\|X\|_X\}$ then
   a. If $n=1$ then as $H=H^1$ we have that H^1 is continous and if $x\in X$ we have $\|H^1(x)\|_X=\|H(x)\|_X\leqslant\|H\|\cdot\|x\|_X=\|H\|^1\cdot\|x\|_X$ proving that $1\in S$
   b. If $n\in S$ then if $x\in X$ we have $\|H^{n+1}(x)\|_X=\|H(H^n(x))\|_X\leqslant\|H\|\cdot\|H^n(x)\|_X\leqslant_{n\in S}\|H\|\cdot(\|H\|^n\cdot\|x\|_X)=\|H\|^{n+1}\cdot\|x\|$ proving that $n+1\in S$
3. **(continuity of $\sum_{i=1}^n H^i$)** We use induction again so let $S=\{n\in\mathbb{N}|\sum_{i=1}^n H^i\text{ is continuous with }\forall x\in X\vDash\|\sum_{i=1}^n H^i(x)\|_X\leqslant(\sum_{i=1}^n\|H\|^i)\}$ then we have:
   a. If $n=1$ then $\forall x\in X$ we have $\|\sum_{i=1}^1 H^i(x)\|_X=\|H^1(x)\|=\|H(x)\|_X\leqslant\|H\|\cdot\|x\|_X=(\sum_{i=1}^1\|H\|^i)\cdot\|x\|_X$ so that $1\in S$.
   b. If $n\in S$ then $\|\sum_{i=1}^{n+1}H^i(x)\|_X=\|H^{n+1}(x)+\sum_{i=1}^n H^i(x)\|_X\leqslant\|H^{n+1}(x)\|_X+\|\sum_{i=1}^n H^i(x)\|_X\leqslant_{n\in S}\|H^{n+1}(x)\|_X+(\sum_{i=1}^n\|H\|^i)\cdot\|x\|_X\leqslant_{(2)}\|H\|^{n+1}\cdot\|x\|_X+(\sum_{i=1}^n\|H\|^i)\cdot\|x\|_X=(\|H\|^{n+1}+\sum_{i=1}^n\|H\|^i)\cdot\|x\|_X=(\sum_{i=1}^{n+1}\|H\|^i)\cdot\|x\|_X$ proving that $n+1\in S$. □

Definition 14.136. *Let $\langle X,\|\|_X\rangle$ be a normed space, if $H\in L(X,Y)$ is such that $\forall x\in X$ we have that $\sum_{i=1}^\infty H^i(x)$ converges then we say that $\sum_{i=1}^\infty H^i$ converges and define $(\sum_{i=1}^\infty H^i)(x)=\sum_{i=1}^\infty H^i(x)$ or $(\sum_{i=1}^\infty H^i)(x)=\lim_{n\to\infty}(\sum_{i=1}^n H^i)(x)$*

Lemma 14.137. *Let $\langle X,\|\|_X\rangle$ be a normed space $L,H\in L(X,Y)$ and $\sum_{i=1}^\infty H^i$ converges then $\forall x\in X$ we have $\lim_{i\to\infty}(L(x)+\sum_{i=1}^n H^i(x))=(L+\sum_{i=1}^\infty H^i)(x)$*

Proof. $(L+\sum_{i=1}^\infty H^i)(x)=L(x)+(\sum_{i=1}^\infty H^i)(x)\underset{\text{definition}}{=}L(x)+\lim_{n\to\infty}(\sum_{i=1}^n H^i(x))=\lim_{n\to\infty}(L(x))+\lim_{n\to\infty}(\sum_{i=1}^n H^i(x))\underset{12.341}{=}\lim_{n\to\infty}(L(x)+\sum_{i=1}^n H^i(x))$ □

Theorem 14.138. *Let $\langle X,\|\|_X\rangle$ be a Banach space and $H\in L(X,X)$ with $\|H\|<1$ then $1_X-H\in\mathcal{G}L(X)$. Further $\sum_{i=1}^\infty H^i$ converges and $(1_X-H)^{-1}=1_X+\sum_{i=1}^\infty H^i$*

Proof. Using the previous lemma (see 14.135) we have that

$$\forall x\in X,n\in\mathbb{N}\text{ we have }\|H^n(x)\|_X\leqslant\|H\|^n\cdot\|x\|_X \tag{14.114}$$

Using the fact that $\|H\|<1$ and 12.403 we have that $\sum_{i=1}^\infty\|H\|^i$ is convergent and $\sum_{i=1}^\infty\|H\|^i=\frac{\|H\|}{1-\|H\|}$, further using 12.369 we have that $\sum_{i=1}^\infty(\|H\|^i\cdot\|x\|_X)$ is convergent to $\frac{\|H\|\cdot\|x\|_X}{1-\|H\|}$. Using the fact that X is a Banach space, 14.114 and 12.381 we have that $\forall x\in X$ that $\sum_{i=1}^\infty H^i(x)$ is convergent and

$$\left\|\sum_{i=1}^\infty H^i(x)\right\|_X\leqslant\frac{\|H\|\cdot\|x\|}{1-\|H\|} \tag{14.115}$$

Also by 12.364 we have that $\lim_{n\rightarrow\infty} H(x)^n = 0$ so that

$$\lim_{n\rightarrow\infty} H(x)^n = 0 \tag{14.116}$$

So we can define the function $(1_X + \sum_{i=1}^{\infty} H^i): X \rightarrow X$ by

$$\forall x \in X \vDash \left(1_X + \sum_{i=1}^{\infty} H^i\right)(x) = 1_X + \sum_{i=1}^{\infty} H^n(x) \tag{14.117}$$

We show now that $(1_X + \sum_{i=1}^{\infty} H^i)$ is linear

$$\begin{aligned}
\left(1_X + \sum_{i=1}^{\infty} H^i\right)(\alpha \cdot x + \beta \cdot y) &= \alpha \cdot x + \beta \cdot y + \sum_{i=1}^{\infty} H^i(\alpha \cdot x + \beta \cdot y) \\
&\underset{14.135}{=} \alpha \cdot x + \beta \cdot y + \sum_{i=1}^{\infty} (\alpha \cdot H^i(x) + \beta \cdot H^i(y)) \\
&\underset{12.369}{=} \alpha \cdot x + \beta \cdot y + \alpha \cdot \sum_{i=1}^{\infty} H^i(x) + \beta \cdot \sum_{i=1}^{\infty} H^i(y) \\
&= \alpha \cdot \left(1_X + \sum_{i=1}^{\infty} H^i\right)(x) + \beta \cdot \left(1_X + \sum_{i=1}^{\infty} H^i\right)(y)
\end{aligned}$$

Next we prove continuity of $(1_X + \sum_{i=1}^{\infty} H^i)$

$$\begin{aligned}
\left\|\left(1_x + \sum_{i=1}^{\infty} H^i\right)(x)\right\|_X &\leqslant \left\|1_x(x) + \sum_{i=1}^{\infty} H^i(x)\right\|_X \\
&\leqslant \|x\|_X + \left\|\sum_{i=1}^{\infty} H^i(x)\right\|_X \\
&\leqslant_{14.115} \|x\|_X + \frac{\|H\|}{1 - \|H\|} \cdot \|x\|_X \\
&= \left(1 + \frac{\|H\|}{1 - \|H\|}\right) \cdot \|x\|_X \\
&= \frac{1}{1 - \|H\|} \cdot \|x\|
\end{aligned}$$

proving continuity of $(1_X + \sum_{i=1}^{\infty} H^i)$ and that

$$\left\|\left(1_X + \sum_{i=1}^{\infty} H^i\right)\right\| \leqslant \frac{1}{1 - \|H\|} \tag{14.118}$$

Next we prove by induction on n that

$$(1_X - H) \circ \left(1_X + \sum_{i=1}^{\infty} H^i\right) = (1_X - H^{n+1}) \tag{14.119}$$

Proof. So let $S=\{n\in\mathbb{N}|(1_X-H)\circ(1_X+\sum_{i=1}^n H^i)=(1_X-H^{n+1})\}$ then we have:

1. If $n=1$ then $\forall x\in X$ we have

$$\begin{array}{rcl}(1_X-H)\left(\left(1_X+\sum_{i=1}^{1} H^i\right)(x)\right) & = & (1_X-H)\left(1_X(x)+\sum_{i=1}^{1} H^i(x)\right)\\ & = & (1_X-H)(x+H(x))\\ & = & x+H(x)-H(x)-H^2(x)\\ & = & (1_X-H^2)(x)\end{array}$$

   proving that $1\in S$.

2. If $n\in S$ then $\forall x\in X$ we have

$$\begin{array}{rcl}(1_X-H)\left(\left(1_X+\sum_{i=1}^{n+1} H^i\right)(x)\right) & = & (1_X-H)\left(\left(1_X+\sum_{i=1}^{n} H^i\right)(x)+H^{n+1}(x)\right)\\ & = & (1_X-H)\left(\left(1_X+\sum_{i=1}^{n} H^i\right)(x)\right)+(1_X-H)(H^{n+1}(x))\\ & \underset{n\in S}{=} & (1_X-H^{n+1})(x)+H^{n+1}(x)-H^{(n+1)+1}(x)\\ & = & (1_X-H^{(n+1)+1})(x)\end{array}$$

   proving that $n+1\in S$.

Using induction we have that $S=\mathbb{N}$ so that 14.118 is valid. □

Next by induction again we prove that

$$\left(1_X+\sum_{i=1}^{n} H^i\right)\circ(1_X-H)=1_X-H^{n+1} \tag{14.120}$$

Proof. So let $S=\{n\in\mathbb{N}|(1_X+\sum_{i=1}^n H^i)\circ(1_X-H)=1_X-H^{n+1}\}$ then we have :

1. If $n=1$ then $\forall x\in X$

$$\begin{array}{rcl}\left(1_X+\sum_{i=1}^{1} H^i\right)((1_X-H)(x)) & = & (x-H(x)+H(x-H(x)))\\ & = & x-H^2(x)\\ & = & (1_X-H^2)(x)\end{array}$$

   proving that $1\in S$.

2. If $n\in S$ then we have $\forall x\in X$ we have

$$\begin{aligned}
\left(1_X+\sum_{i=1}^{n+1} H^i\right)((1_X-H)(x)) &= 1_X(x-H(x))+\sum_{i=1}^{n} H^i(x-H(x))+\\
&\quad H^{n+1}(x-H(x))\\
&= \left(1_X+\sum_{i=1}^{n}\right)((1_X-H)(x))+\\
&\quad H^{n+1}(x)-H^{(n+1)+1}(x)\\
&\underset{n\in S}{=} (1_X-H^{n+1})(x)+H^{n+1}(x)-\\
&\quad H^{(n+1)+1}(x)\\
&= x-H^{(n+1)+1}(x)\\
&= (1_X-H^{(n+1)+1})(x)
\end{aligned}$$

proving that $n+1\in S$.

Using induction we have that $S=\mathbb{N}$ so that 14.120 is indeed valid. □

Now if $\varepsilon>0$, $x\in X$ then by 14.116 there exists a $N\in\mathbb{N}$ such that if $n\geqslant N$ then $\|H^{n+1}(x)\|_X<\varepsilon$ so that using 14.119 and 14.120 we have $\|(1_X-H)((1_X-\sum_{i=1}^n H^i)(x))-1_X(x)\|_X=\|-H^{n+1}(x)\|_X=\|H^{n+1}(x)\|_X<\varepsilon$ and $\|(1_X+\sum_{i=1}^n H^i)((1_X-H)(x))-1_X\|_X=\|-H^{n+1}(x)\|_X=\|H^{n+1}(x)\|_X<\varepsilon$ proving that

$$\forall x\in X\vDash \lim_{n\rightarrow\infty}\left[\left(1_X+\sum_{i=1}^n H^i\right)((1_X-H)(x))\right]=1_X(x) \tag{14.121}$$

Using the above and 14.137 we have that $(1_X+\sum_{i=1}^\infty H^i)((1_X-H)(x))=1_X(x)$ so that

$$\left(1_X+\sum_{i=1}^\infty H^i\right)\circ(1_X-H)=1_X \tag{14.122}$$

$$\forall x\in X\vDash \lim_{n\rightarrow\infty}\left[(1_X-H)\left(\left(1_X+\sum_{i=1}^n H^i\right)(x)\right)\right]=1_X(x) \tag{14.123}$$

As (1_X-H) is continuous we have by 12.323 and the above that $1_X(x)=\lim_{n\rightarrow\infty}[(1_X-H)((1_X+\sum_{i=1}^n H^i)(x))]=(1_X-H)(\lim_{n\rightarrow\infty}(1_X+\sum_{i=1}^n H^i)(x))=(1_X-H)(\lim_{n\rightarrow\infty}(1_X(x)+\sum_{i=1}^n H^i(x)))\underset{14.137}{=}(1_X-H)((1_X+\sum_{i=1}^\infty H^i)(x))$ proving that

$$1_X=(1_X-H)\circ\left(1_X+\sum_{i=1}^\infty H^i\right) \tag{14.124}$$

Finally using 14.122 and 14.124 we have that $(1_X-H)^{-1}=1_X+\sum_{i=1}^\infty H^i$. Proving that 1_X-H is a bijection and as $1_X, H\in L(X,X)$ we have also that 1_X-H is linear and continuous thus a isomorphism and by 14.130 it is toplinear. □

Corollary 14.139. *Let $\langle X, \|\|_X\rangle$ be a Banach space, $H \in L(X, X)$ with $\|H\| < 1$ then $1_X + H \in \mathcal{GL}(X)$ with $(1_X + H)^{-1} = \sum_{i=1}^{\infty} (-1)^i \cdot H^i$*

Proof. This follows from the previous theorem by using $-H$ as $\|-H\| = \|H\| < 1$ □

Theorem 14.140. *Let $\langle X, \|\|\rangle$ be a Banach space then we have*

1. *$\mathcal{GL}(X)$ is a open set in $L(X, X)$*
2. *The function $\tau: \mathcal{GL}(X) \rightarrow \mathcal{GL}(X)$ defined by $L \rightarrow \tau(L) = L^{-1}$ is differentiable of class C^∞ and $D\tau(L) \in L(\mathcal{GL}(X), L(\mathcal{GL}(X), \mathcal{GL}(X)))$ is defined by $D\tau(L)(H) = -L^{-1} \circ H \circ L^{-1}$*

Proof.

1. Let $L \in \mathcal{GL}(X)$ then for any $H \in L(X, X)$ we can write

$$\begin{aligned} H &= L + (H - L) \\ &= L \circ 1_X + (L \circ L^{-1}) \circ (H - L) \\ &= L \circ (1_X + L^{-1} \circ (H - L)) \end{aligned}$$

giving

$$H = L \circ (1_X + L^{-1} \circ (H - L)) \tag{14.125}$$

Now according to 12.183 we have $\|L^{-1} \circ (H - L)\| \leqslant \|L^{-1}\| \cdot \|H - L\|$, take then $\delta = \max(1, \|L^{-1}\|) > 0$ then if $\|H - L\| < \frac{1}{\delta}$ we have $\|L^{-1} \circ (H - L)\| < \|L^{-1}\| \cdot \frac{1}{\delta} \leqslant 1$ and using the previous corollary we have then that $(1_X + L^{-1} \circ (H - L)) \in \mathcal{GL}(X)$ and as $L \in \mathcal{GL}(X)$ we have by 14.132 that $H = L \circ (1_X + L^{-1} \circ (H - L)) \in \mathcal{GL}(X)$. So if $H \in B_{\|\|}\left(L, \frac{1}{\delta}\right)$ then $\|H - L\| < \frac{1}{\delta}$ and thus $H \in \mathcal{GL}(X) \Rightarrow B_{\|\|}\left(L, \frac{1}{\delta}\right)$ proving that $\mathcal{GL}(X)$ is open.

2. Next we prove that τ is differentiable at every $L \in \mathcal{GL}(X)$

   a. First we prove that τ is differentiable at $1_X \in \mathcal{GL}(X)$ with $\forall H \in L(X, X)$ that $D\tau(1_X)(H) = -1_X^{-1} \circ H \circ 1_X^{-1}$. Note that if $\|H\| < 1$ then using the previous corollary we have $1_X + H \in \mathcal{GL}(X) \Rightarrow 1_X \in B_{\|\|}(1_X, 1) \subseteq \mathcal{GL}(X)$ and $(1_X + H) = 1_X + \sum_{i=1}^{\infty} (-1)^i \cdot H^i$. So we have

$$\begin{aligned} \tau(1_X + H) - \tau(1_X) &= (1_X + H)^{-1} - 1_X \\ &= 1_X + \sum_{i=1}^{\infty} (-1)^i \cdot H^i - 1_X \\ &= \sum_{i=1}^{\infty} (-1)^i \cdot H^i \\ &\underset{12.367}{=} -H + \sum_{i=2}^{\infty} (-1)^i \cdot H^i \end{aligned}$$

   giving

$$\tau(1_X + H) - \tau(1_X) = -H + \sum_{i=2}^{\infty} (-1)^i \cdot H^i \tag{14.126}$$

As $1_X+H\in\mathcal{G}L(X)$ we have $1_X+\sum_{i=1}^{\infty}(-1)^i\cdot H^i\in\mathcal{G}L(X)\subseteq L(X,X)\underset{1_X\in L(X,X)}{\Rightarrow}\sum_{i=1}^{\infty}(-1)^i\cdot H^i\in L(X,X)\underset{H\in L(X,X)}{\Rightarrow}\sum_{i=2}^{\infty}(-1)^i\cdot H^i=H+\sum_{i=1}^{\infty}(-1)^i\cdot H^i\in L(X,X)$ proving that

$$\sum_{i=2}^{\infty}(-1)^i\cdot H^i\in L(X,X) \tag{14.127}$$

Then as $C_{X,0}\in L(X,X)$ we have using 14.127 that $\varepsilon(H)=\begin{cases}C_{X,0}\text{ if }H=0\\ \frac{1}{\|H\|}\cdot\sum_{i=2}^{\infty}(-1)^n\cdot H^i, H\neq 0\end{cases}\in L(X,X)$. This defines then a function ε

$$\varepsilon\colon B_{\|\|}(0,1)=(B_{\|\|}(1_X,1))_{1_X}\rightarrow L(X,X),\quad H\rightarrow\varepsilon(H)=\begin{cases}C_{X,0}\text{ if }H=0\\ \frac{1}{\|H\|}\cdot\sum_{i=2}^{\infty}(-1)^n\cdot H^i, H\neq 0\end{cases} \tag{14.128}$$

So using 14.126 we have

$$\tau(1_X+H)-\tau(1_X)+(-H)=\varepsilon(H)\cdot\|H\| \tag{14.129}$$

As $\|H\|<1$ we can use 12.403 and 12.367 we have that $\|H\|+\sum_{i=2}^{\infty}\|H\|^i=\frac{\|H\|}{1-\|H\|}\Rightarrow\sum_{i=2}^{\infty}\|H\|^i=\frac{\|H\|-\|H\|\cdot(1-\|H\|)}{1-\|H\|}=\frac{\|H\|^2}{1-\|H\|}\underset{12.369}{\Rightarrow}\sum_{i=2}^{\infty}\|H\|^i\cdot\|x\|_X=\frac{\|H\|^2\cdot\|x\|_X}{1-\|H\|}$. Now if $0<\|H\|<1$, $x\in X$ then as $\forall i\in\mathbb{N}$ we have that $\|(-1)^n\cdot H^n(x)\|_X\leqslant\|H^n(x)\|_X\leqslant\|H\|^n\cdot\|x\|$ we can use 12.381 to prove that $\|\sum_{i=2}^{\infty}(-1)^i\cdot H^i(x)\|_X\leqslant\frac{\|H\|^2}{1-\|H\|}\cdot\|x\|_X$ so that $\|\varepsilon(H)(x)\|_X=\left\|\frac{1}{\|H\|}\cdot\sum_{i=2}^{\infty}(-1)^i\cdot H^i(x)\right\|_X\leqslant\frac{\|H\|}{1-\|H\|}\cdot\|x\|_X$ giving that

$$\|\varepsilon(H)\|\leqslant\frac{\|H\|}{1-\|H\|} \tag{14.130}$$

Now given $\xi>0$ choose $0<\|H\|<\min\left(\frac{\xi}{2},\frac{1}{2}\right)$ then $\|\varepsilon(H)-\varepsilon(0)\|=\|\varepsilon(H)-0\|=\|\varepsilon(H)\|<\frac{\frac{\xi}{2}}{1-\frac{1}{2}}=\xi$ proving that ε is continuous at $0\in L(X,X)$ and thus a ε-mapping. Proving by 14.13, 14.129 that if $-*\colon L(X,X)\rightarrow L(X,X)$ is defined by $-*(H)=-H$ (which is trivially linear and continuous) we have that τ is differentiable at 1_X with $D\tau(1_X)=-*$ so that $D\tau(1_X)(H)=-H=-(1_X^{-1}\circ H\circ 1_X^{-1})$.

b. Next we prove that τ is differentiable at a arbitrary $L\in\mathcal{G}L(X)$ with $D\tau(L)(H)=-L^{-1}\circ H\circ L^{-1}$. So take $\delta=\min\left(\frac{1}{1+\|L^{-1}\|},1\right)$ then if $\|H\|<\delta$ we have $\|H\circ L^{-1}\|\leqslant\|H\|\cdot\|L^{-1}\|<\delta\cdot\|L^{-1}\|=\frac{\|L^{-1}\|}{1+\|L^{-1}\|}\leqslant 1$ proving by the previous corollary that

$$\text{if } H\in B_{\|\|}(0,\delta)\Rightarrow 1_X+H\circ L^{-1}\in\mathcal{G}L(X). \tag{14.131}$$

As $L\in\mathcal{G}L(X)$ we have by 14.131 that

$$L+H=(1_X+H\circ L^{-1})\circ L\in\mathcal{G}L(X) \tag{14.132}$$

or

$$L \in B_{\|\|}(L, \delta) \subseteq \mathcal{G}L(X) \tag{14.133}$$

So as also $\delta < 1$ we have if $H \in B_{\|\|}(0, \delta) = (B_{\|\|}(L, \delta))_L \Rightarrow \|H \circ L^{-1}\| \leqslant \|H\| \cdot \|L^{-1}\| < \frac{1}{1 + \|L^{-1}\|} \cdot \|L^{-1}\| \leqslant 1 \Rightarrow \varepsilon(H \circ L^{-1})$ is defined, then

$$\begin{aligned}
\tau(L+H) - \tau(H) &= (L+H)^{-1} - L^{-1} \\
&= ((1_X + H \circ L^{-1}) \circ L) - L^{-1} \\
&= L^{-1} \circ (1_X + H \circ L^{-1}) - L^{-1} \\
&= L^{-1} \circ ((1_X + H \circ L^{-1})^{-1} - 1_X) \\
&= L^{-1} \circ (\tau(1_X + H \circ L^{-1}) - \tau(1_X)) \\
&= L^{-1} \circ (D\tau(1_X)(H \circ L^{-1}) + \varepsilon(H \circ L^{-1}) \cdot \|H \circ \\
&\quad\ L^{-1}\|) \\
&= L^{-1} \circ (-(H \circ L^{-1}) + \varepsilon(H \circ L^{-1}) \cdot \|H \circ L^{-1}\|) \\
&= -L^{-1} \circ H \circ L^{-1} + L^{-1} \circ \varepsilon(H \circ L^{-1}) \cdot \|H \circ L^{-1}\|
\end{aligned}$$

giving

$$\tau(L+H) - \tau(H) - (-L^{-1} \circ H \circ L^{-1}) = L^{-1} \circ \varepsilon(H \circ L^{-1}) \cdot \|H \circ L^{-1}\| \tag{14.134}$$

Given $H \in B_{\|\|}(0, \delta) = (B_{\|\|}(L, \delta))_L$ define then $\varepsilon'(H) = \begin{cases} 0 \text{ if } H = 0 \\ L^{-1} \circ \varepsilon(H \circ L^{-1}) \cdot \frac{\|H \circ L^{-1}\|}{\|H\|} \text{ if } H \neq 0 \end{cases}$, note that as $\varepsilon(H \circ L^{-1}) \in L(X, X)$ and $L^{-1} \in L(X, X)$ we can define a function ε' as follows

ε': $B_{\|\|}(0, \delta) = (B_{\|\|}(L, \delta))_L \rightarrow L(X, X)$ by $\varepsilon'(H) =$

$$\begin{cases} 0 \text{ if } H = 0 \\ L^{-1} \circ \varepsilon(H \circ L^{-1}) \cdot \frac{\|H \circ L^{-1}\|}{\|H\|} \text{ if } H \neq 0 \end{cases} \tag{14.135}$$

Define now $(-L^{-1} \circ * \circ L^{-1})$: $L(X, X) \rightarrow L(X, X)$ by $H \rightarrow (-L^{-1} \circ * \circ L^{-1})(H) = -L^{-1} \circ H \circ L^{-1}$. So using the above and 14.134 we have that

$$\tau(L+H) - \tau(H) - (-L^{-1} \circ * \circ L)(H) = \varepsilon'(H) \cdot \|H\| \tag{14.136}$$

Next we prove that $(-L^{-1} \circ * \circ L^{-1})$ is linear and continuous.

Proof.

i. **(linearity)** $(-L^{-1} \circ * \circ L^{-1})(\alpha \cdot H + \beta \cdot H') = -L^{-1} \circ (\alpha \cdot H + \beta \cdot H') \circ L^{-1} = \alpha \cdot [-L^{-1} \circ H \circ L^{-1}] + \beta \cdot [-L^{-1} \circ H' \circ L^{-1}] = \alpha \cdot (-L^{-1} \circ * \circ L^{-1})(H) + \beta \cdot (-L^{-1} \circ * \circ L^{-1})(H')$

ii. **(continuity)** For $H \in L(X, X)$ we have $\|(-L^{-1} \circ * \circ L^{-1})(H)\|_X = \|-L^{-1} \circ H \circ L^{-1}\|_X = \|L^{-1} \circ H \circ L^{-1}\| \leqslant \|L^{-1}\|^2 \cdot \|H\|$ proving continuity

□

Next we prove that ε' is continue at 0 and thus a ε-mapping.

Proof. If $\xi > 0$ then as ε is continue at 0 there exists a $\delta_1 > 0$ such that if $\|H\| < \delta_1$ we have $\|\varepsilon(H)\| < \frac{\xi}{\|L^{-1}\|^2 + 1}$. So if $\delta_2 = \frac{\delta_1}{\|L^{-1}\| + 1}$ then if $0 < \|H\| < \delta_2 \Rightarrow \|H \circ L^{-1}\| \leqslant \|H\| \cdot \|L^{-1}\| < \delta_2 \cdot \|L^{-1}\| \leqslant \delta_1 \Rightarrow$ $\|\varepsilon(H \circ L^{-1})\| < \frac{\xi}{\|L^{-1}\|^2 + 1}$ so that $\|\varepsilon'(H) - \varepsilon'(0)\| = \|\varepsilon'(H) - 0\| =$ $\|\varepsilon'(H)\| = \left\| L^{-1} \circ \varepsilon(H \circ L^{-1}) \cdot \frac{\|H \circ L^{-1}\|}{\|H\|} \right\| \leqslant \frac{\|H \circ L^{-1}\|}{\|H\|} \cdot \|L^{-1}\| \cdot \|\varepsilon(H \circ$ $L^{-1})\| \leqslant \frac{\|H\| \cdot \|L^{-1}\|}{\|H\|} \cdot \|L^{-1}\| \cdot \|\varepsilon(H \circ L^{-1})\| = \|L^{-1}\|^2 \cdot \|\varepsilon(H \circ L^{-1})\| <$ $\|L^{-1}\|^2 \cdot \frac{\xi}{\|L^{-1}\|^2 + 1} < \xi$. □

The above together with 14.136 proves that τ is differentiable at L with $D\tau(L) = (-L^{-1} \circ * \circ L^{-1})$ so that $D\tau(L)(H) = -L^{-1} \circ H \circ L^{-1}$.

3. Next to prove that τ is C^∞ define first the following functions:

   a. Define $\chi: \mathcal{GL}(X) \rightarrow \mathcal{GL}(X) \times \mathcal{GL}(X)$ defined by $L \rightarrow \chi(L) = (L, L)$ then we have that $\pi_1 \circ \chi = 1_{\mathcal{GL}(X)}$ is C^∞ and $\pi_2 \circ \chi = 1_{\mathcal{GL}(X)}$ is C^∞ so that by 14.66 χ is C^∞.

   b. Define $\rho: \mathcal{GL}(X) \times \mathcal{GL}(X) \rightarrow L(L(X, X), L(X, X))$ by $(L, L') \rightarrow (-L \circ * \circ L')$ where $(-L \circ * \circ L')(H) = -L \circ H \circ L'$ then we have:

      i. **(bi linearity)** $\forall H \in L(X, X)$ we have $\rho(\alpha \cdot L_1 + \beta \cdot L_2, L)(H) = -(\alpha \cdot L_1 + \beta \cdot L_2) \circ H \circ L = \alpha \cdot (-L_1 \circ H \circ L) + \beta \cdot (-L_2 \circ H \circ L) = \alpha \cdot \rho\,(L_1, L)(H) + \beta \cdot \rho(L_2, L)(H) = (\alpha \cdot \rho\,(L_1, L) + \beta \cdot \rho(L_2, L))(H)$ proving that $\rho(\alpha \cdot L_1 + \beta \cdot L_2, L) = \alpha \cdot \rho(L_1) + \beta \cdot \rho(L_2)$. $\forall H \in L(X, X)$ we have $\rho(L, \alpha \cdot L_1 + \beta \cdot L_2)(H) = -L \circ H \circ (\alpha \cdot L_1 + \beta \cdot L_2) \underset{L, H \text{is linear} \Rightarrow L \circ H \text{is linear}}{=} \alpha \cdot (-L \circ H \circ L_1) + \beta \cdot (-L \circ H \circ L_2) = \alpha \cdot \rho(L, L_1)(H) + \beta \cdot \rho(L, L_2)(H) = (\alpha \cdot \rho(L, L_1) + \beta \cdot \rho(L, L_2))(H)$ proving that $\rho(L, \alpha \cdot L_1 + \beta \cdot L_2) = \alpha \cdot \rho(L, L_1) + \beta \cdot \rho(L, L_2)$.

      ii. **(continuity)** If $(L_1, L_2) \in \mathcal{GL}(X) \times \mathcal{GL}(X, X)$ and $H \in L(X, X)$ then we have $\|\rho(L_1, L_2)(H)\| = \|-L_1 \circ H \circ L_2\| \leqslant \|L_1\| \cdot \|L_2\|$ proving that $\|\rho(L_1, L_2)\| \leqslant \|L_1\| \cdot \|L_2\| = 1 \cdot \|L_1\| \cdot \|L_2\|$ proving togheter with bi linearity (thus multilinearity) that ρ is continuous with $\|\rho\| \leqslant 1$.

      Using 14.63 we have then that ρ is C^∞

   c. Take now $\rho \circ \chi \circ \tau$ and $L \in \mathcal{GL}(X)$ then $(\rho \circ \chi \circ \tau)(L) = \rho(\chi(\tau(L))) = \rho(\chi(L^{-1})) = \rho(L^{-1}, L^{-1}) = (-L^{-1} \circ * \circ L^{-1}) = D\tau(L)$ thus $\rho \circ \chi \circ \tau = D\tau$. We proceed now by induction so let $S = \{n \in \mathbb{N} | \tau \text{ is } C^n\}$ then we have:

      i. If $n = 1$, as τ is differentiable at every $L \in \mathcal{GL}(X)$ we have that τ is continuous at every $L \in \mathcal{GL}(X)$ and thus as χ, ρ are C^∞ and thus continuous we have that $D\tau = \rho \circ \chi \circ \tau$ is continuous, proving that $\tau = C^1$ so that $1 \in S$.

      ii. Assume now that $n \in S$ then τ is C^n so that by the Chain Rule and the fact that $\chi, \rho - C^\infty$ so that $D\tau = \rho \circ \chi \circ \tau$ is C^n proving by 14.54 that τ is C^{n+1}.

By mathematical induction we have that $S = \mathbb{N}$ and thus $\tau = C^{\infty}$ $\square$

Definition 14.141. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces $U \subseteq X$ a open set, $V \subseteq Y$ a open set then a bijective function $f: U \rightarrow V$ is a diffeomorphism (of class C^r, $r \in \mathbb{N}_0$) if and only if $f: U \rightarrow V$, $f^{-1}: V \rightarrow U$ are differentiable on U, V [of class C^r, $r \in \mathbb{N}_0$]*

Example 14.142. Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces then a homeomorphism $f: X \rightarrow Y$ is a diffeomorphism of class C^0.

Note 14.143. Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces $U \subseteq X$ a open set, $V \subseteq Y$ a open and a function $f: U \rightarrow V$ then using 2.40 we have that f is a diffeomorphism [of class C^r] if and only if f is differentiable [of class C^r] and there exists a function $g: V \rightarrow U$ such that g is differentiable [of class C^r] and $f \circ g = 1_V$ and $g \circ f = 1_U$.

Theorem 14.144. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces, $U \subseteq X$ a open set, $V \subseteq Y$ a open set and $f: U \rightarrow V$ a diffeomorphism and $g = f^{-1}: V \rightarrow U$ then if we take $x \in U$ with $y = f(x)$ we have $\mathrm{Dg}(y) \circ \mathrm{Df}(x) = 1_X$ and $\mathrm{Df}(x) \circ \mathrm{Dg}(y) = 1_Y$ proving that $\mathrm{Df}(x)$ and $\mathrm{Dg}(y)$ are toplinear isomorphisms.*

Proof. Let $y = f(x) \Rightarrow g(y) = x$ then we have

1. From $f \circ g = 1_V$ and $1_Y = D1_V(x) = D(f \circ g)(y) = \mathrm{Df}(g(y)) \circ \mathrm{Dg}(y) = \mathrm{Df}(x) \circ \mathrm{Dg}(y)$
2. From $g \circ f = 1_U$ and $1_X = D1_U(x) = D(g \circ f)(x) = \mathrm{Dg}(f(x)) \circ \mathrm{Df}(x) = \mathrm{Dg}(y) \circ \mathrm{Df}(x)$ $\square$

Theorem 14.145. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$, $\langle Z, \|\|_Z\rangle$ be normed spaces, $U \subseteq X$ a open set, $V \subseteq Y$ a open set, $W \subseteq Z$ a open set and $f: U \rightarrow V$ a diffeomorphism [of class C^r], $g: V \rightarrow W$ a diffeomorphism [of class C^r] then $f \circ g: U \rightarrow W$ is a diffeomorphism [of class C^r]*

Proof. As $f \circ g$ is a bijection with $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$ (see 2.46) we have by the Chain Rule (see 14.68). $\square$

Definition 14.146. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces, $U \subseteq X$ then a function $f: U \rightarrow Y$ is a local diffeomorphism [of class C^r] at $x \in U$ if there exists a open set $U(x) \subseteq U, V(f(x)) \subseteq Y$ with $x \in U(x) \subseteq U$, $f(x) \subseteq V(f(x))$ and $f_{|U(x)}$: $U(x) \rightarrow V(f(x))$ is a diffeomorphism [of class C^r]*

We state now the Inverse function theorem

Theorem 14.147. (The Inverse Function Theorem) *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|_Y\|\rangle$ be Banach spaces over $\mathbb{R}$, $U \subseteq X$ an open set, $f: U \rightarrow Y$ a differentiable function of class C^r, $r \in \mathbb{N}$. Assume that for a $x_0 \in U$ we have that $\mathrm{Df}(x_0): X \rightarrow Y$ is a isomorphism then f is a local diffeomorphism of class C^r at x_0. In other words there exists open sets $U(x_0) \subseteq U \subseteq X$, $V(f(x_0)) \subseteq Y$ with $x_0 \in U(x_0)$, $f(x_0) \subseteq V(f(x_0))$ such that $f_{|U(x_0)}: U(x_0) \rightarrow V(f(x_0))$ is a diffeomorphism.*

Proof. The proof is quite long and is done by proving several propositions. First note that by the definition of the differential (see 14.6) we have that $\mathrm{Df}(x_0)$ is a linear continuous isomorphism so by 14.130 and the fact that X, Y are Banach spaces we have that

$$\mathrm{Df}(x_0)\,\text{is}\,a\,\text{toplinear isomorphism} \tag{14.137}$$

First we prove a more specific version of the the Inverse Function Theorem.

Proposition 14.148. *Let $\langle X, \|\|_X\rangle$ be a Banach space over $\mathbb{R}$, $0 \in U \subseteq X$ an open set $f: U \rightarrow X$ a differentiable function of class $C^r, r \in \mathbb{N}$ such that $\mathrm{Df}(0) = 1_X$ and $f(0) = 0$ then f is a is a local diffeomorphism at 0.*

Proof. Define $T: U \rightarrow X$ by $x \rightarrow T(x) = 1_X(x) - f(x)$ then $T(0) = 0$ and T is of class C^r (as it is the sum of a C^∞ and a C^r function (see 14.58)) so

$$D^1 T\,\text{is continuous on}\,U \tag{14.138}$$

Next $D^1 T(0) \underset{14.58}{=} D^1 1_X(0) - D^1 f(0) = 1_X - 1_X = 0$ [$0 \in L(X, X)$ the neutral element in $L(X, Y)$] giving

$$D^1 T(0) = 0 \tag{14.139}$$

As $D^1 T$ is continuous at 0 there exists a ρ' with $0 \in B_{\|\|_X}(0, \rho') \subseteq U$ such that $\forall x \in B_{\|\|_X}(0, \rho')$ we have that $D^1 T(x) \in B_{\|\|}\left(0, \frac{1}{2}\right) \Rightarrow \|D^1 T(x)\| < \frac{1}{2}$. Take then $\rho = \frac{\rho'}{2}$ then $\bar{B}_{\|\|_X}(0, \rho) \subseteq B_{\|\|}(0, \rho') \subseteq U$ so that we have

$$\text{if}\,\|x\|_X \leqslant \rho\,(\text{or}\,x \in B_{\|\|_X}(0, \rho))\,\text{then}\,\|D^1 T(x)\| < \frac{1}{2} < 1 \tag{14.140}$$

Now as $D^1 T(x) = 1_X - D^1 f(x) \Rightarrow D^1 f(x) = 1_X - D^1 T(x)$ we can apply 14.138 to get

$$\text{If}\,\|x\|_X \leqslant \rho\,(\text{or}\,x \in B_{\|\|_X}(0, \rho))\,\text{then}\,D f(x)\,\text{is}\,a\,\text{toplinear isomorphism} \tag{14.141}$$

Now using the fact that $\bar{B}_{\|\|_X}(0, \rho)$ is convex (see 14.101) and the fundamental theorem of calculus (see 14.102) together with 14.140 we find that if $x \in \bar{B}_{\|\|_X}(0, \delta)$ then $\|T(x)\|_Y = \|T(x) - T(0)\|_X \leqslant \frac{1}{2} \cdot \|x - 0\|_X = \frac{1}{2} \cdot \|x\|_X$ giving

$$\text{If}\,x \in \bar{B}_{\|\|_X}(0, \rho)\,\text{then}\,\|T(x)\|_X \leqslant \frac{1}{2} \cdot \|x\|_X \tag{14.142}$$

Next we prove the following proposition.

Proposition 14.149. *Let $0 < \sigma \leqslant \rho$ then $\forall y \in \bar{B}_{\|\|_Y}\left(0, \frac{\sigma}{2}\right)$ there exists a unique determined $x \in \bar{B}_{\|\|_X}(0, \sigma)$ such that $f(x) = y$*

Proof. First note that $\bar{B}_{\|\|_X}(0, \sigma), \bar{B}_{\|\|_X}\left(0, \frac{\sigma}{2}\right)$ are closed (see 12.56) and thus by 12.336 and the completeness of X is complete. Define now $T_y = C_{\bar{B}_{\|\|_X}(0,\sigma),y} + T: \bar{B}_{\|\|_X}(0, \sigma) \rightarrow X$ by $x \rightarrow T_y(x) = y + T(x) = y + 1_X(x) - f(x) = y + x - f(x)$. If $x \in \bar{B}_{\|\|_X}(0, \sigma) \subseteq \bar{B}_{\|\|_X}(0, \rho)$ so that $\|T_y(x)\|_Y = \|y + T(x)\|_X \leqslant \|y\|_X + \|T(x)\|_X \leqslant_{y \in \bar{B}_{\|\|_X}(0, \frac{\sigma}{2})} \frac{\sigma}{2} + \|T(x)\|_X \leqslant_{14.142} \frac{\sigma}{2} + \frac{x}{2} \leqslant \frac{\sigma}{2} + \frac{\sigma}{2} = \sigma$ giving

$$x \in \bar{B}_{\|\|_X}(0, \sigma) \Rightarrow \|T_y(x)\|_X \leqslant \sigma \tag{14.143}$$

Note that T_y is C^r as it is the sum of the C^∞ function $C_{\bar{B}_{\|\|_X}(0,\sigma)}$ and the C^r function T with $D^1 T_y = 0 + D^1 T = D^1 T$ so that by 14.140 we have $\forall x \in \bar{B}_{\|\|_X}(0,\sigma) \subseteq \bar{B}_{\|\|_X}(0, \rho)$ that $D^1 T_y(x) < \frac{1}{2}$ so that using 14.102 we have

$$\forall x_1, x_2 \in \bar{B}_{\|\|_X}(0,\sigma) \text{ we have } \|T_y(x_1) - T_y(x_2)\|_X \leqslant \frac{1}{2} \cdot \|x_1 - x_2\|_X \tag{14.144}$$

proving that T_y is a contraction (see 12.400). Using the Banach Fixed Point theorem (see 12.404) and the fact that $\bar{B}_{\|\|_X}(0,\sigma)$ is complete there exists a unique $x \in \bar{B}_{\|\|_X}(0,\sigma)$ such that $T_y(x) = x \Leftrightarrow x = y + x - f(x) \Leftrightarrow y = f(x)$ proving the proposition. □

Define now

$$\text{Given } \chi > 0 \text{ define } 0 \in U_\chi = B_{\|\|_X}(0, \chi) \bigcap f^{-1}\left(B_{\|\|_X}\left(0, \frac{\chi}{2}\right)\right) a \text{ open set (as } f \text{ is cont.)} \tag{14.145}$$

We can then prove the following proposition

Proposition 14.150. *There exists a χ with $0 < \chi < \rho$ such that $f_{|U_\chi}: U_\chi \rightarrow B_{\|\|_X}\left(0, \frac{\chi}{2}\right)$ is a bijective function.*

Proof. Take $\chi = \frac{\rho}{2}$ then $0 < \chi < \rho$ and we have

1. **(Surjectivity)** Using the previous proposition (14.149) we have if $y \in B_{\|\|_X}\left(0, \frac{\chi}{2}\right) \subseteq \bar{B}_{\|\|_X}\left(0, \frac{\chi}{2}\right)$ that there exists a unique $x \in \bar{B}_{\|\|_X}(0, \chi)$ such that $f(x) = y$. We prove now by contradiction that $\|x\|_X \neq \chi$. [So assume that $\|x\|_X = \mathcal{X}$ take then $\varepsilon = \frac{\chi}{2} - \|y\|_X > 0$ then by the continuity of f (as f is C^r) there exists a $\tau' > 0$ such that if $\|x - x'\|_X < \tau'$ then $\|f(x) - f(x')\|_X < \varepsilon$. Take now $\tau = \min\left(\tau', \frac{\rho}{4}\right)$ and $x_1 = x \cdot \left(1 + \frac{\tau}{\chi}\right)$ then we have $\|x_1 - x\|_X = \frac{\tau}{\chi} \cdot \|x\|_X = \tau < \tau'$ so that $\|f(x_1)\|_X \leqslant \|f(x) - f(x_1)\|_X + \|f(x)\|_X < \frac{\chi}{2} - \|y\|_X + \|y\|_X = \frac{\chi}{2} \Rightarrow f(x_1) \in \bar{B}_{\|\|_X}\left(0, \frac{\chi}{2}\right)$ which as $0 < \chi < \rho$ means that by the previous proposition there exists a $x_2 \in \bar{B}_{\|\|_X}(0, \chi)$ such that $f(x_1) = f(x_2)$. As $\|x_1\|_X = \|x\|_X \cdot \left(1 + \frac{\tau}{\chi}\right) = \chi + \tau < \frac{\rho}{2} + \frac{\rho}{4} = \frac{3}{4} \cdot \rho < \rho$ giving

$$x_1 \in \bar{B}_{\|\|_X}(0, \rho) \tag{14.146}$$

As $x_2 \in \bar{B}_{\|\|_X}\left(0, \frac{\chi}{2}\right) = \bar{B}_{\|\|_X}\left(0, \frac{\rho}{2}\right) \subseteq \bar{B}_{\|\|_X}(0, \rho)$ giving

$$x_2 \in \bar{B}_{\|\|_X}(0, \rho) \tag{14.147}$$

Also $f(x_1) = f(x_2) \in \bar{B}_{\|\|_X}\left(0, \frac{\chi}{2}\right) = \bar{B}_{\|\|_X}\left(0, \frac{\rho}{2}\right)$ proving

$$f(x_1) = f(x_2) \in \bar{B}_{\|\|_X}\left(0, \frac{\rho}{2}\right) \tag{14.148}$$

Using the fact that $0<\rho\leqslant\rho$, 14.147, 14.148 the previous proposition (especially the uniqueness) we reach the conclusion $x_1=x_2$ but then $\|x_2\|_X\leqslant\chi<\chi\cdot\left(1+\frac{\tau}{\chi}\right)=\left\|x\cdot\left(1+\frac{\tau}{\chi}\right)\right\|_X=\|x_1\|_X\Rightarrow\|x_2\|_X<\|x_1\|_X$ contradicting $x_1=x_2$.]. So as $\|x\|_X\neq\chi$ and $x\in\bar{B}_{\|\|_X}(0,\chi)$ we must have that $x\in B_{\|\|_X}(0,\chi)\underset{f(x)=y\in B_{\|\|_X}(0,\frac{\chi}{2})}{\Rightarrow}x\in U\chi$ proving surjectivity.

2. **(Injectivity)** Assume that there exists a $x_1,x_2\in U_\chi\subseteq B_{\|\|_X}(0,\chi)\subseteq\bar{B}_{\|\|_X}(0,\chi)$ which as $f(x_1),f(x_2)\in f(U_\chi)\subseteq B_{\|\|_X}\left(0,\frac{\chi}{2}\right)$ which by the previous proposition we have $x_1=x_2$. □

Now by the above proposition we have a χ with $0<\chi<\rho$ such that $f_{|U_X}\colon U_X\to B_{\|\|_X}\left(0,\frac{\chi}{2}\right)$ is a bijection. By the bijectivity the function $g=(f_{|U_\chi})^{-1}\colon B_{\|\|_X}\left(0,\frac{\chi}{2}\right)\to U_\chi$ will then exists.

Proposition 14.151. *g is continuous on $B_{\|\|_X}\left(0,\frac{\chi}{2}\right)$*

Proof. If $x\in U_\chi\subseteq\bar{B}_{\|\|_X}(0,\rho)$ a convex set then $x=x-f(x)+f(x)=T(x)+f(x)$ so if $x_1,x_2\in U_\chi$ then

$$\begin{aligned}\|x_1-x_2\|_X &= \|T(x_1)+f(x_1)-T(x_2)-f(x_2)\|_X\\ &\leqslant \|T(x_1)-T(x_2)\|_X+\|f(x_1)-f(x_2)\|_X\end{aligned}$$

so that by using 14.140, the convexity of $\bar{B}_{\|\|_X}(0,\rho)$ and the mean value theorem (see 14.102) we have $\|x_1-x_2\|_X\leqslant\frac{1}{2}\cdot\|x_1-x_2\|_X+\|f(x_1)-f(x_2)\|_X$ so that we have

$$\|x_1-x_2\|_X\leqslant 2\cdot\|f(x_1)-f(x_2)\|_X \tag{14.149}$$

If now $y_1,y_2\in B_{\|\|_X}\left(0,\frac{\chi}{2}\right)$ then $g(y_1),g(y_2)\in U_\chi$ and we have using 14.149 that $\|g(y_1)-g(y_2)\|_X\leqslant 2\cdot\|f(g(y_1))-f(g(y_2))\|_X=2\cdot\|y_1-y_2\|_X$ giving

$$\|g(y_1)-g(y_2)\|_X\leqslant 2\cdot\|y_1-y_2\| \tag{14.150}$$

So if $\varepsilon>0$ and $y\in B_{\|\|_X}\left(0,\frac{\chi}{2}\right)$ then for all $y'\in B_{\|\|_X}\left(0,\frac{\chi}{2}\right)$ with $\|y-y'\|_X<\frac{\varepsilon}{2}$ we have $\|g(y)-g(y')\|_X\leqslant 2\cdot\frac{\varepsilon}{2}=\varepsilon$ proving that g is continuous on $B_{\|\|_X}\left(0,\frac{\chi}{2}\right)$. □

Next we prove that g is differentiable.

Proposition 14.152. *g is differentiable on $B_{\|\|_X}\left(0,\frac{\chi}{2}\right)$ and $\mathrm{Dg}(y)=(\mathrm{Df}(g(y)))^{-1}$*

Proof. Now if $y\in B_{\|\|_X}\left(0,\frac{\chi}{2}\right)$ we have that $g(y)\in U_\chi\underset{14.145}{\Rightarrow}g(y)\in B_{\|\|_X}(0,\chi)\subseteq B_{\|\|_X}(0,\rho)\subseteq U$ we have by 14.141 that $\mathrm{Df}(g(y))$ is a toplinear isomorphism so that $\mathrm{Df}(g(y))^{-1}$ is continuous and linear. Take now $h\in\left(B_{\|\|_X}\left(0,\frac{\chi}{2}\right)\right)_y\Rightarrow y+h\in B_{\|\|_X}\left(0,\frac{\chi}{2}\right)$ let

$$k=g(y+h)-g(y) \tag{14.151}$$

then $k+g(y)=g(y+h)\in g\left(B_{\|\|_X}\left(0,\frac{\chi}{2}\right)\right)=U_\chi\subseteq B_{\|\|_X}(0,\rho)\subseteq U\Rightarrow k\in(B_{\|\|_X}(0,\rho))_{g(y)}\subseteq U_{g(y)}$ and as f is differentiable at $g(y)$ there exists a ε-mapping $\zeta\colon U_{g(y)}\to X$ such that

$$f(g(y)+k)-f(g(y))-\mathrm{Df}(g(y))(k)=\zeta(k)\cdot\|k\|_X \tag{14.152}$$

So using $f(g(y)+k)=f(g(y+h))=y+h$ and $f(g(y))=y$ and thus $f(g(y)+k)-f(g(y))=h$ we find that

$$h=\mathrm{Df}(g(y))(k)+\zeta(k)\cdot\|k\|_X \tag{14.153}$$

Applying $\mathrm{Df}(g(y))^{-1}$ to both sides of 14.153 and using linearity we have $\mathrm{Df}(g(y))^{-1}(h)=\mathrm{Df}(g(y))^{-1}(\mathrm{Df}(g(y))(k)+\zeta(k)\cdot\|k\|_X)=\mathrm{Df}(g(y))^{-1}(\mathrm{Df}(g(y))(k))+\mathrm{Df}(g(y))^{-1}(\zeta(k))\cdot\|k\|_X$ giving $\mathrm{Df}(g(y))^{-1}(h)=k+\mathrm{Df}(g(y))^{-1}(\zeta(k))\cdot\|k\|_X=g(y+h)-g(y)+\mathrm{Df}(g(y))^{-1}(\zeta(k))\cdot\|k\|_X$. So we have

$$\mathrm{Df}(g(y))^{-1}(h)=g(y+h)-g(y)+\mathrm{Df}(g(y))^{-1}(\zeta(k))\cdot\|k\|_X \tag{14.154}$$

giving

$$g(y+h)-g(y)-\mathrm{Df}(g(y))^{-1}(h)=-\mathrm{Df}(g(y))^{-1}(\zeta(g(y+h)-g(y)))\cdot\|g(y+h)-g(y)\|_X \tag{14.155}$$

Define now given $\varepsilon\colon \left(B_{\|\|_X}\left(0,\ \frac{\chi}{2}\right)\right)_y \to X$ where $h\to\varepsilon(h)=\begin{cases}0 \text{ if } h=0\\ -\mathrm{Df}(g(y))^{-1}(\zeta(g(y+h)-g(y)))\cdot\frac{\|g(y+h)-g(y)\|_X}{\|h\|_X} \text{ if } h\neq 0\end{cases}$ then we have that

$$g(y+h)-g(y)-\mathrm{Df}(g(y))^{-1}(h)=\varepsilon(h)\cdot\|h\|_X \tag{14.156}$$

Note that in the proof of the previous proposition (see 14.150) we have $\|g(y+h)-g(y)\|_X\leqslant 2\cdot\|y+h-y\|_X=2||h|$ hence if $h\neq 0$ then

$$\frac{\|g(y+h)-g(y)\|_X}{\|h\|_X}\leqslant 2 \tag{14.157}$$

So let $\gamma<0$ then by continuity of ζ at 0 there exists a $\delta'>0$ such that if $\|z\|_X<\delta'$ then $\|\zeta(z)\|_X<\gamma$. By continuity of g at y there exists a $\delta''>0$ such that if $\|h\|_X<\delta''$ then $\|g(y+h)-g(y)\|_X<\delta'$ and thus we have

$$\text{Given}\gamma>0\,\text{there exists a}\,\delta''>0\,\text{such that}\,\|h\|_X<\delta''\,\text{then}\,\|\zeta(h)\|_X<\gamma \tag{14.158}$$

If now $\gamma>0$ find then by 14.158 a $\delta''>0$ such that if $\|h\|_X<\delta''$ then $\|\zeta(h)\|_X<\frac{\gamma}{2\cdot(\|\mathrm{Df}(g(y))^{-1}\|+1)}$ so thate $\|\varepsilon(h)-\varepsilon(0)\|_X=\|\varepsilon(h)\|_X=\frac{\|g(y+h)-g(y)\|_X}{\|h\|_X}\cdot\|-\mathrm{Df}(g(y))^{-1}(\zeta(g(y+h)-g(y)))\|_X\leqslant 2\cdot\|\mathrm{Df}(g(y))^{-1}(\zeta(g(y+h)-g(y)))\|_X\leqslant 2\cdot\|\mathrm{Df}(g(y))^{-1}\|\cdot\|\zeta(h)\|_X<\gamma$ proving that ε is continuous at 0 and thus a ε-mapping. This together with the linearity ,continuity of $\mathrm{Df}(g(y))^{-1}$ and 14.156 we have that g is differentiable at y with differential $\mathrm{Df}(g(y))^{-1}=\mathrm{Dg}(y)$. □

If we take then $\tau\colon L(X,X)\to L(X,X)$ defined by $L\to\tau(L)=L^{-1}$ (see 14.140) then we have if $y\in B_{\|\|_X}\left(0,\frac{\chi}{2}\right)$ that $(\tau\circ\mathrm{Df}\circ g)(y)=\tau(\mathrm{Df}(g(y)))=\mathrm{Df}(g(y))^{-1}=\mathrm{Dg}(y)$ proving that

$$\tau\circ\mathrm{Df}\circ g=\mathrm{Dg} \tag{14.159}$$

Let's use induction to prove that g is C^r, so let $S=\{n\in\mathbb{N}|\text{if } n\leqslant r \text{ then } g \text{ is } C^n\}$ then we have

1. If $n=1$ then using 14.140 we have that τ is C^∞ and thus τ is continuous, by assumption Df is continuous and we have proved that g is continuous (see 14.151) so we must have that Dg is continuous, proving that g is C^1 and thus that $1\in S$.

2. If $n \in S$ then if $n+1 \leqslant r$ we have that g is C^n, τ is C^∞ and f is C^r so that $\mathrm{Dg} = \tau \circ \mathrm{Df} \circ g$ is C^n so that g is C^{n+1} proving that $n+1 \in S$

Using induction we have then that $S = \mathbb{N}$ so as $r \in \mathbb{N} = S$ and $r \leqslant r$ we have g is C^r. This proves that f is a local C^r diffeomorphism at 0 and thus proves this proposition (see 14.148).

□

Next we must extend proposition 14.148 to the more general case.

Proposition 14.153. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|_Y\|\rangle$ be Banach spaces over $\mathbb{R}$, $U \subseteq X$ an open set, $f: U \rightarrow Y$ a differentiable function of class C^r, $r \in \mathbb{N}$. Assume that for a $0 \in U$ we have that $f(0) = 0$ and that $Df(0): X \rightarrow Y$ is a isomorphism then f is a local diffeomorphism of class C^r at x_0.*

Proof. As $\mathrm{Df}(0)$ is continuous and linear and a isomorphism we have by 14.130 that $\mathrm{Df}(0)$ is toplinear so that $\mathrm{Df}(0)^{-1}: Y \rightarrow X$ is defined, linear and continuous and thus by 14.61 is C^∞. Define then $h = \mathrm{Df}(0)^{-1} \circ f: U \rightarrow X$ which is is differentiable on U of class C^r. Then $h(0) = \mathrm{Df}(0)^{-1}(f(0)) = \mathrm{Df}(0)^{-1}(0) \underset{\text{linearity}}{=} 0$, also $\mathrm{Dh}(0) = D(\mathrm{Df}(0)^{-1})(f(0)) \circ \mathrm{Df}(0) = \mathrm{Df}(0)^{-1} \circ \mathrm{Df}(0) = 1_X$. So by 14.148 there exists $U(0), V'(0)$ open in X with $0 \in U(0), 0 \in V'(0)$ such that $h_{|U(0)}: U(0) \rightarrow V'(0)$ is a diffeomorphism of class C^r. As $\mathrm{Df}(0)$ is continuous and linear it is C^∞ so also $\mathrm{Df}(0)_{|V'(0)}: V'(0) \rightarrow \mathrm{Df}(0)(V'(0)) = V(0)$ is C^∞ and is isomorphic thus toplinear. as $0 \in V'(0)$ and $\mathrm{Df}(0)_{|V'(0)}(0) = \mathrm{Df}(0)(0) \underset{\text{linearity}}{=} 0 \Rightarrow 0 \in V(0)$, as $\mathrm{Df}(0)_{|V'(0)}$ is toplinear (and thus open) we have that $0 \in V(0)$ is open. Take now $g = \mathrm{Df}(0)_{|V'(0)} \circ h_{|U(0)}$: $U(0) \rightarrow V(0)$ then as it is a composition of diffeomorphism of class C^r it is a diffeomorphism of class C^r (see 14.145). If now $x \in U(0)$ then we have $g(x) = \mathrm{Df}(0)_{|V'(0)}(h_{|U(0)|}(0)) = \mathrm{Df}(0)(\mathrm{Df}(0)^{-1}(f(x))) = f(x)$ so that $f_{|U(0)} = g$ and thus $f_{|U(0)}: U(0) \rightarrow V(0)$ is a diffeomorphism of class C^r □

Finally we extend the above proposition to the general case where $f(0) \neq 0$ so lets repeat the Inverse Function Theorem

Proposition 14.154. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|_Y\|\rangle$ be Banach spaces over $\mathbb{R}$, $U \subseteq X$ an open set, $f: U \rightarrow Y$ a differentiable function of class C^r, $r \in \mathbb{N}$. Assume that for a $x_0 \in U$ we have that $Df(x_0): X \rightarrow Y$ is a isomorphism then f is a local diffeomorphism of class C^r at x_0.*

Proof. First given $y \in Y$ define $t_y^Y: Y \rightarrow Y$ defined by $x \rightarrow x + y$ which is trivially a C^∞ diffeomorphism with $\mathrm{Dt}_y^Y = 1_Y$ and $(t_y^Y)^{-1} = t_{-y}^Y$, also $t_y^X: X \rightarrow X$ defined by $x \rightarrow x + y$ is a C^∞ diffeomorphism with $\mathrm{Dt}_y^X = 1_X$ and $(t_y^X)^{-1} = t_{-y}^X$. Consider then $h = t_{-f(x_0)}^Y \circ f \circ t_{x_0}^X: t_{x_0}^X(U) \rightarrow Y$ then $h: t_{x_0}^X(U) \rightarrow Y$ is a C^r differentiable function of class C^r with $h(0) = t_{-f(x_0)}^Y(f(t_{x_0}^X(0))) = t_{-f(x_0)}^Y(f(x_0 + 0)) = t_{-f(x_0)}^Y(f(x_0)) = f(x_0) - f(x_0) = 0$. So we can use the previous proposition (see 14.153) to find open sets $U'(0) \subseteq X, V'(0) \subseteq Y$ with $0 \in U'(0), 0 \in V'(0)$ such that $h_{|U(0)}: U'(0) \rightarrow V'(0)$ is a diffeomorphism of class C^r. Take then $g = t_{f(x_0)}^Y \circ h_{|U'(0)} \circ t_{-x_0}^X: t_{-x_0}^X(t_{x_0}^X(U'(0))) = U(0) \rightarrow t_{f(x_0)}^Y(V'(0)) = V(0)$ which is a diffeomorphism of class C^r. If $x \in U(0)$ then $g(x) = t_{f(x_0)}^Y(h_{|U'(0)}(t_{-x_0}^X(x))) = t_{f(x_0)}^Y(h(x - x_0)) = t_{f(x_0)}^Y(t_{-f(x_0)}^Y(f(t_{x_0}^X(x - x_0)))) =$

$f(x)$ proving that $g_{|U(0)} = f_{|U(0)}$ so that $f_{|U(0)}: U(0) \rightarrow V(0)$ is a diffeomorphism of class C^r as also $0 \in U'(0) \Rightarrow x_0 = t^X_{x_0}(0) \in t^X_{x_0}(U'(0)) = U(0)$ and $0 \in V'(0) \Rightarrow f(x_0) = t^Y_{f(x_0)}(0) \in t^Y_{f(x_0)}(V'(0)) = V(0)$ proving that f is a local diffeomorphism of class C^r. $\square$

The above proposition proves finally the Inverse Function Theorem $\square$

Corollary 14.155. *Given $\langle\mathbb{K}, ||\rangle$ (where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$), a normed space, $U \subseteq \mathbb{K}$ a open set, $x \in U$ a open set, $V \subseteq \mathbb{K}$ and a function $f: U \rightarrow V$ such that $\forall x \in U$ we have that $f'(x)$ exists and is continuous and for $x_0 \in U$ we have that $f'(x_0) \neq 0$ then there exists open sets U', V' with $x_0 \in U' \subseteq U$ and $f(x_0) \in V' \subseteq V$ such that $f_{|U'}: U' \rightarrow V'$ is a bijection and $(f^{-1}_{|U'}): V' \rightarrow U'$ has a derivative $(f^{-1}_{|U'})'(x)$ at every $x \in V'$ such that $(f^{-1}_{|U'})'(x) \cdot f'(f^{-1}(x)) = 1$.*

Proof. As the existance of a derivative at a point implies differentiability at this point (see 14.17). Further we have that $Df: U \rightarrow L(\mathbb{K}; \mathbb{K})$ is defined by $x \rightarrow Df(x)$ where $Df(x)(t) = t \cdot f'(x)$. From the continuity of f' it follows that given a $\varepsilon > 0$ there exists a $\delta > 0$ such that $\forall y \in U$ with $|y - x| < \varepsilon$ we have that $\|f'(x) - f'(y)\| < \varepsilon$. Further $\forall t \in \mathbb{K}$ we have $\|(Df(x) - Df(y))(t)\| = \|Df(x)(t) = Df(y)(t)\| = \|t \cdot f'(x) - t \cdot f'(y)\| = |t| \cdot \|f'(x) - f'(y)\| < t \cdot \varepsilon$ which proves that Df is continuous on U. So

$$f \text{ is } C^1 \tag{14.160}$$

Further at x_0 we have that $Df(x_0): \mathbb{K} \rightarrow \mathbb{K}$ is defined by $t \rightarrow f'(x_0) \cdot t$ if we define g: $\mathbb{K} \rightarrow \mathbb{K}$ by $t \rightarrow \frac{t}{f'(x_0)}$ then we have that $(Df(x_0) \circ g)(t) = \left(\frac{t}{f'(x_0)}\right) \cdot f'(x_0) = t = 1_{\mathbb{K}}(t)$ and $(g \circ Df(x_0))(t) = (f'(x_0) \cdot t) \cdot \frac{1}{f'(x_0)} = t = 1_{\mathbb{K}}(t)$ proving that

$$Df(x_0) \text{ is a bijection and } (Df(x_0))^{-1} \tag{14.161}$$

Using 14.160 and 14.161 we can apply the Inverse Function Theorem (see 14.147) then f is a local diffeomorphism of class $C>$ Hence there exists a U', V' open with $x_0 \in U' \subseteq U$ and $f(x_0) \in V' \subseteq V$ such that $f_{|U'}: U' \rightarrow V'$ is a diffeomorphism. Hence $(f^{-1})_{|V'} = (f_{|U'})^{-1}: V' \rightarrow U'$ is differentiable on V' and thus by 14.17 $(f^{-1})_{|V'}$ has a derivative at every point on V'. As $f_{|U'} \circ f^{-1}_{|U'} = 1_{U'}$ we have that $\forall x \in V'$ that $1 = (1_{|U'})'(x) \underset{14.29}{=} (f_{|U'})'(f^{-1}_{|U'}(x)) \cdot (f^{-1}_{|U'})'(x) \underset{14.18}{=} f'(f'_{|U'}{}^{-1}(x)) \cdot (f^{-1}_{|U'})'(x) = f'(f^{-1}(x)) \cdot (f^{-1}_{|U'})'(x)$ proving that

$$f'(f^{-1}(x)) \cdot (f^{-1}_{|U'})'(x) = 1$$ $\square$

A import conclusion of the Inverse Function Theorem is the below Implicit Function Theorem.

Lemma 14.156. *Let X, Y, Z be vector spaces over the field $\mathbb{K}$ and $L \in \operatorname{Hom}(X, Z)$ and $H \in \operatorname{Hom}(Y, Z)$ such that $H^{-1} \in \operatorname{Hom}(Z, Y)$ exists then $T: X \times Y \rightarrow X \times Z$ defined by $(x, y) \rightarrow (x, L(x) + H(y))$ has a inverse $S: X \times Z \rightarrow X \times Y$ defined by $(x, z) \rightarrow (x, -H^{-1}(L(x)) + H^{-1}(z))$ so that T is a we have bijection and thus a isomorphism.*

Proof. If $(x,z)\in X\times Z$ then $T(S(x,z))=T(x,-H^{-1}(L(x))+H^{-1}(z))=(x,L(x)+H(-H^{-1}(L(x))+H^{-1}(z)))\underset{\text{H is linear}}{=}(x,L(x)-H(H^{-1}(L(x)))+H(H^{-1}(z)))=(x,L(x)-L(x)+z)=(x,z)$. Also if $(x,y)\in X\times Y$ then $S(T(x,y))=S(x,L(x)+H(y))=(x,-H^{-1}(L(x))+H^{-1}(L(x)+H(y)))\underset{H^{-1}\text{ is also linear}}{=}(x,-H^{-1}(L(x))+H^{-1}(L(x))+H^{-1}(H(y)))=(x,y)$ $\square$

Theorem 14.157. (Implicit Function Theorem) *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ and $\langle Z,\|\|_Z\rangle$ be Banach spaces, $U\subseteq X$, $V\subseteq Y$ be open sets and let $f:U\times V\rightarrow Z$ be a differentiable function of class $C^r, r\in\mathbb{N}$. Assume that $\exists x_0\in U$, $\exists y_0\in V$ such that $D_2f(x_0,y_0):Y\rightarrow Z$ is a isomorphism then there exists a open set $W_0\subseteq U\times Z$ and a open set $V_0\subseteq V$ with $(x_0,f(x_0,y_0))\in W_0)$ such that there exists a unique C^r map $g:W_0\rightarrow V_0$ such that $g(x_0,f(x_0,y_0))=y_0$ and $\forall x,z\in W_0$ we have $f(x,g(x,z))=z$, furthermore we have that*

- $D_1g(x,z)=-[D_2f(x,g(x,z))]^{-1}\circ D_1f(x,g(x,z))$
- $D_2g(x,z)=[D_2f(x,g(x,z))]^{-1}$

Proof. *Define $\varphi:U\times V\rightarrow X\times Z$ by $(x,y)\rightarrow\varphi(x,y)=(x,f(x,y))$ then as $\pi_1^{X\times Z}\circ\varphi=(\pi_1^{X\times Y})_{|U\times V}$ a restriction of a continuous linear function which is C^∞ with differential $D^1(\pi_1^{X\times Z}\varphi)(x,y)=\pi_1^{X\times Y}$ and $\pi_2^{X\times Z}\circ\varphi=f$ we have that $\pi_2^{X\times Z}\circ\varphi$ is C^r and by 14.123 we have that $D^1(\pi_2^{X\times Z}\circ\varphi)(x,y)=D_1^1f(x)\circ\pi_1^{X\times Y}+D_2^1f(x)\circ\pi_2^{X\times Y}$. Using 14.66 and 14.65 we have that φ is C^r with $\pi_1\circ D^1\varphi(x,y)=\pi_1^{X\times Y}$ and $\pi_2\circ D^1\varphi(x,y)=D_1^1f(x)\circ\pi_1^{X\times Y}+D_2^1f(x)\circ\pi_2$. So at $(x_0,y_0)\in U\times V$ we have that $D^1f(x_0,y_0)(x,y)=(x,D_1^1f(x_0,y_0)(x)+D_2^1f(x_0,y_0)(y))$ Using the above lemma and the fact that $D_2f(x_0,y_0)$ is a isomorphism we have that $D^1\varphi(x_0,y_0)$ is a isomorphism. Using the Inverse Function Theorem (see 14.147), so there exists open sets $U_{(x_0,y_0)}\subseteq U\times V, V_{\varphi(x_0,y_0)}\subseteq X\times Z$ with $(x_0,y_0)\in U_{x_0}$ and $(x_0,f(x_0,y_0))=\varphi(x_0,y_0)\in V_{\varphi(x_0,y_0)}$ such that*

$$\varphi_{|U(x_0,y_0)}:U_{(x_0,y_0)}\rightarrow V_{\varphi(x_0,y_0)}\text{ is a local diffeomorphisme of class }C^r \tag{14.162}$$

where

$$W_0\underset{\text{def}}{=}V_{\varphi(x_0,y)}=\varphi(U_{(x_0,y_0)})\subseteq U\times Z\text{ and }\varphi_{|U_{(x_0,y_0)}}(x_0,y_0)=(x_0,f(x_0,y_0)) \tag{14.163}$$

So $(\varphi_{|U_{(x_0,y_0)}})^{-1}:V_{\varphi(x_0,y_0)}\rightarrow U_{(x_0,y_0)}$ exists and is C^r. So if $(x,z)\in V_{\varphi(x_0,y_0)}$ then

$$\begin{aligned}(x,z) &= \varphi_{|U_{(x_0,y_0)}}((\varphi_{|U_{(x_0,y_0)}})^{-1}(x,z))\\ &= \varphi((\varphi_{|U_{(x_0,y_0)}})^{-1}(x,z))\\ &= \varphi(\pi_1^{X\times Y}((\varphi_{|U_{(x_0,y_0)}})^{-1}(x,z)),\pi_2^{X\times Y}((\varphi_{|U_{(x_0,y_0)}})^{-1}(x,z)))\\ &= (\pi_1^{X\times Y}((\varphi_{|U_{(x_0,y_0)}})^{-1}(x,z)),f(\pi_1^{X\times Y}((\varphi_{|U_{(x_0,y_0)}})^{-1}(x,z)),\\ &\quad\pi_2^{X\times Y}((\varphi_{|U_{(x_0,y_0)}})^{-1}(x,z))))\end{aligned}$$

so that $x=\pi_1^{X\times Y}((\varphi_{|U_{(x_0,y_0)}})^{-1}(x,z))$ and thus we have

$$\begin{aligned}(x,z) &= (x,f(x,\pi_2^{X\times Y}((\varphi_{|U_{(x_0,y_0)}})^{-1}(x,z))))\\ &\Rightarrow\\ z &= f(x,\pi_2^{X\times Y}((\varphi_{|U_{(x_0,y_0)}})^{-1}(x,z)))\end{aligned}$$

so if we take

$$g: W_0 \rightarrow V \text{ by } (x, z) \rightarrow g(z) = \pi_2^{X \times Y}((\varphi_{|U_{(x_0, y_0)}})^{-1}(x, z)) \in \pi_2^{X \times Y}(U_{(x_0, y_0)}) = V_0 \subseteq V \tag{14.164}$$

where

$$V_0 = \pi_2^{X \times Y}(U_{(x_0, y_0)}) \text{ is open because of } 12.144$$

we have

$$\forall (x, z) \in W_0 = V_{\varphi(x_0, y_0)} \text{ we have } f(x, g(x, z)) = z \tag{14.165}$$

As from 14.163 we have $\varphi_{|u_{(X_0, Y_0)}}(x_0, y_0) = (x_0, f(x_0, y_0))$ *it follows that* $(\varphi_{|U_{(x_0, y_0)}})^{-1}(x_0, f(x_0, y_0)) = (x_0, y_0) \Rightarrow y_0 = \pi_2((\varphi_{|U_{(x_0, y_0)}})^{-1}(x_0, f(x_0, y_0))) = g(x_0, f(x_0, y_0))$ *proving that*

$$g(x_0, f(x_0, y_0)) = y_0 \tag{14.166}$$

Next as $\pi_2^{X \times Y}$ *is* C^∞ *(being continuous and linear) and* $(\varphi_{|U_{(x_0, y_0)}})^{-1}$ *is* C^r *we have that*

$$g \text{ is } C^r \tag{14.167}$$

and we have proved the first assertment of the theorem.

To prove uniqueness assume, that there is a $h: W_0 \rightarrow V_0 = \pi_2^{X \times Y}(U_{(x_0, y_0)})$ *with* $\forall (x, z) \in W_0$ $f(x, h(x, z)) = z$ *so that* $\exists (x_1, z_1)$ *with* $g(x_1, z_1) \neq h(x_1, z_1)$ *and derivate a contradiction. As* $(x_1, z_1) \in W_0 = V_{\varphi(x_0, y_0)} \underset{14.162}{=} \varphi_{|U(x_0, y_0)}(U_{(x_0, y_0)})$ *there exists a* $(x_2, y_2) \in U_{(x_0, y_0)}$ *such that* $(x_1, z_1) = \varphi_{|U_{(x_0, y_0)}}(x_2, y_2) = \varphi(x_2, y_2) = (x_2, f(x_2, y_2)) \Rightarrow x_1 = x_2 \in \pi_1^{X \times Y}(U_{(x_0, y_0)})$ *and as* $h(x_1, z_1), g(x_1, z_1) \in V_0 = \pi_2^{X \times Y}(U_{(x, y_0)})$ *proving that* $(x_1, g(x_1, z_1)), (x_2, h(x_2, z_2)) \in U_{(x_0, y_0)}$ *so that* $\varphi_{|U_{(x_0, y_0)}}(x_1, g(x_1, z_1)) = \varphi(x_1, g(x_1, z_1)) = (x_1, f(x_1, g(x_1, z_1))) = (x_1, z_1) = (x_1, f(x_1, h(x_1, z_1))) = \varphi(x_1, h(x_1, z_1)) = \varphi_{|U_{(x_0, y_0)}}(x_1, h(x_1, z_1)) \underset{\varphi_{|U_{(x_0, y_0)}} \text{is injective}}{\Rightarrow} (x_1, g(x_1, z_1)) = (x_1, h(x_1, z_1)) \Rightarrow g(x_1, z_1) = h(x_1, z_1)$ *contradicting* $g(x_1, z_1) \neq h(x_1, z_1)$.

Now to prove the differentials of g note that if $z \in W_0$ *is fixed then we can define*

$$\psi: W_0 \rightarrow U \times V \text{ by } (x, z) \rightarrow \psi(x, z) = (x, g(x, z)) \text{ (see 14.163 and 14.164)} \tag{14.168}$$

if we define then also

$$f \circ \psi: W_0 \rightarrow Z \ (x, z) = \omega(x, z) = f(x, g(x, z)) = z = \pi_2(x, z) \Rightarrow f \circ \psi = \pi_2 \tag{14.169}$$

As $z = \pi_2(x, z) = \pi_2((*, z)(x)) = (\pi_2 \circ (*, z))(x) = C_{W_0, z}(x)$ *we have that* $D_1(f \circ \psi)(x, z) = D((f \circ \psi) \circ (*, z))(x) = D(\pi_2 \circ (*, z))(x) = D(C_{W_0, z})(x) = C_{X, 0}$ *so that*

$$D_1(f \circ \psi)(x, z) = C_{X, 0} \tag{14.170}$$

Note that $(\psi \circ (*, z))(h) = \psi(h, z) = (h, g(h, z)) = (\pi_1(h, z), g(h, z)) = (\pi_1(\star, z)(h), g((*, z)(h))) = ((\pi_1 \circ (*, z))(h), (g \circ (\star, z))(h)) = (\pi_1 \circ (*, z), g \circ (*, z))(h) \Rightarrow \psi \circ (*, z) = (\pi_1 \circ (*, z), g \circ (*, z))$ *proving that*

$$\pi_1 \circ (\psi \circ (*, z)) = \pi_1 \circ (*, z) = 1_{W_0} \text{ and } \pi_2 \circ (\psi \circ (*, z)) = g \circ (*, z) \tag{14.171}$$

Using 14.124 *we have that* $D_1(f\circ\psi)(x,z)=Df(\psi(x.z))\circ D_1\psi(x,z)=Df(\psi(x,z))\circ D(\psi\circ(*,z))(x) \underset{14.65}{=} Df(\psi(x,z))\circ(D(\pi_1\circ\psi\circ(*,z))(x), D(\pi_2\circ\psi\circ(*,z))(x)) \underset{14.171}{=} Df(\psi(x,z))\circ(D(1_{W_0})(x), D(g\circ(*,z))(x)) = \mathrm{Df}(\psi(x,z))\circ(1_X, D_1g(x,z)) \underset{14.123}{=} D_1f(\psi(x,z))\circ\pi_1\circ(1_X, D_1g(x,z)) + D_2f(\psi(x,z))\circ\pi_2\circ(1_X, D_1g(x,z)) = D_1f(\psi(x,z)) + D_2f(\psi(x,z))\circ D_1g(x,z)$ *proving that*

$$D_1(f\circ\psi)(x,z)=D_1f(\psi(x,z))+D_2f(\psi(x,z))\circ D_1g(x,z) \tag{14.172}$$

Using 14.170 *and* 14.172 *we have that* $D_2f(\psi(x,z))\circ D_1g(x,z)=-D_1f(\psi(x,z))\Rightarrow D_1g(x,z)=-D_2f(\psi(x,z))^{-1}\circ D_1f(\psi(x,z))$ *giving after applying* $\psi(x,z)=(x,g(x,z))$

$$D_1g(x,z)=-D_2f(x,g(x,z))^{-1}\circ D_1f(x,g(x,z)) \tag{14.173}$$

Now $z=\pi_2(x,z)=\pi_2((x,*)(z))=(\pi_2\circ(x,*))(z)=1_{\pi_1(W_0)}(z)$ *so that*

$$\pi_2\circ(x,*)=1_{\pi_1(W_0)} \tag{14.174}$$

So we have that $D_2(f\circ\psi)(x,z)=D_2(\pi_2)(x,z) \underset{14.30}{=} D(\pi_2\circ(x,*))(z)=D(1_{\pi_1(w_0)})=1_X$ *proving that*

$$D_2(f\circ\psi)(x,z)=1_X \tag{14.175}$$

Note that $(\psi\circ(x,*))(h)=\psi(x,h)=(x,g(x,h))=(C_{W_0,x}(h), g((x,*)(h)))=(C_{W_0,x}(h),(g\circ(x,*))(h))=(C_{W_0,x}, g\circ(x,*))(h)$ *proving that*

$$\pi_1\circ(\psi\circ(x,*))=C_{W_0,x}\,\text{and}\,\pi_2\circ(\psi\circ(x,*))=g\circ(x,*) \tag{14.176}$$

Using 14.124 *we have that* $D_2(f\circ\psi)(x,z)=Df(\psi(x,z))\circ D_2\psi(x,z)=Df(\psi(x,z))\circ D(\psi\circ(x,*))(z) \underset{14.65}{=} Df(\psi(x,z))\circ(D(\pi_1\circ\psi\circ(x,*))(z), D(\pi_2\circ\psi\circ(x,*))(z)) \underset{14.176}{=} Df(\psi(x,z))\circ(D(C_{W_0,x})(z), D(g\circ(x,*))(z)) = Df(\psi(x,z))\circ(C_{X,0}, D_2g(x,z)) = D_1f(\psi(x,z))\circ\pi_1\circ(C_{X,0}, D_2g(x,z)) + D_2f(\psi(x,z))\circ\pi_2\circ(C_{X,0}, D_2g(x,z)) = D_2f(\psi(x,z))\circ D_2g(x,z)$ *proving that*

$$D_2(f\circ\psi)(x,z)=D_2f(\psi(x,z))\circ D_2g(x,z) \tag{14.177}$$

Using 14.175 *and* 14.177 *we have then that* $1_X=D_2f(\psi(x,z))\circ D_2g(x,z)\Rightarrow D_2g(x,z)=(D_2f(\psi(x,z)))^{-1}=(D_2f(x,g(x,z)))^{-1}$ *proving*

$$D_2g(x,z)=(D_2f(x,g(x,z)))^{-1} \tag{14.178}$$

Finally 14.173 *and* 14.177 *proves the final part of the theorem.* □

A trivial corrolation of the implicit theorem is

Theorem 14.158. *Let* $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ *and* $\langle Z,\|\|_Z\rangle$ *be Banach spaces,* $U\subseteq X$, $V\subseteq Y$ *be open sets and let* $f:U\times V\rightarrow Z$ *be a differentiable function of class* C^r, $r\in\mathbb{N}$. *Assume that* $\exists x_0\in U$, $\exists y_0\in V$ *such that* $D_2f(x_0,y_0):Y\rightarrow Z$ *is a isomorphism. Then there exists a open set* W *with* $x_0\in W\subseteq U$ *and a function* $g:W\rightarrow V$ *such that* $g(x_0)=y_0$ *and* $\forall x\in W$ *we have* $f(x,g(x))=f(x_0,z_0)$. *Further* $Dg(x)=-(D_2f(x,g(x)))^{-1}\circ D_1f(x,g(x))$

Proof. *Using the Implicit Function Theorem there exists a open set* $W_0 \subseteq U \times Z$ *with* $(x_0, f(x_0, y_0)) \in W_0$ *and a* C^r *function* $g': W_0 \to V$ *such that* $\forall (x, z) \in W_0$ *we have* $f(x, g'(x, z)) = z$ *and* $g'(x_0, f(x_0, y_0)) = y_0$*. Define now* $g: \pi_1(W_0) = W \to V$ *defined by* $x \to g(x) = g'(x, f(x_0, y_0))$ *then we have that*

$$g(x_0) = g'(x_0, f(x_0, y_0)) = y_0$$

and

$$f(x, g(x)) = f(x, g'(x, f(x_0, y_0))) = f(x_0, y_0)$$

as $\pi_1(W_0) = W$ *is open (see 12.144) we have that* $g: W \to V$ *is the function we search for. Finally if* $x \in W$ *then* $(g' \circ (*, f(x_0, y_0)))(x) = g'((*, f(x_0, y_0))(x)) = g'(x, f(x_0, y_0)) = g(x)$ *proving that* $g' \circ (*, f(x_0, y_0)) = g$ *so that* $-[D_2 f(x, g(x))]^{-1} \circ D_1 f(x, g(x)) = -[D_2 f(x, g'(x, f(x_0, y_0)))]^{-1} \circ D_1 f(x, g'(x, f(x_0, y_0))) \underset{\text{Implicit Function Theorem}}{=} D_1 g'(x, f(x_0, y_0)) = D(g' \circ (*, f(x_0, y_0)))(x) = Dg(x)$ *proving that*

$$Dg(x) = -[D_2 f(x, g(x))]^{-1} \circ D_1 f(x, g(x)) \qquad \square$$

Chapter 17

The extended real numbers

17.1 Extented reals

Lemma 17.1. *There exists at least two sets that are not a element of* $\mathbb{R}$

Proof. Using the definition of the real numbers (see 8.1) we have that $\emptyset \notin \mathbb{R}$ and $\mathbb{Q} \notin \mathbb{R}$ $\square$

Using the above lemma and the fact that a set is by definition a element the following definition of extended reals make sense

Definition 17.2. *The set* $\bar{R} = \mathbb{R} \bigcup \{\infty, -\infty\}$ *where* $\infty, -\infty \notin \mathbb{R}$ *and* $\infty \neq -\infty$

We extend the order relation $\leqslant_{\mathbb{R}}$ on $\mathbb{R}$ to a order relation on $\bar{\mathbb{R}}$ that will make $\langle \bar{\mathbb{R}}, \leqslant \rangle$ a fully ordered set. By doing so we will show that as a benefit every non empty set in $\bar{\mathbb{R}}$ has a supremum and a infinum.

Definition 17.3. *Given the fully ordered set* $\langle \mathbb{R}, \leqslant_{\mathbb{R}} \rangle$ *(see 9.27) define* $\leqslant \subseteq \bar{\mathbb{R}} \times \bar{\mathbb{R}}$ *by*

1. $-\infty \leqslant \infty$
2. $-\infty \leqslant -\infty$
3. $\infty \leqslant \infty$
4. $\forall x \in \mathbb{R}$ *we have* $-\infty \leqslant x$ *and* $x \leqslant \infty$
5. $\forall x, y \in \mathbb{R}$ *with* $x \leqslant_{\mathbb{R}} y$ *we have* $x \leqslant y$

Note 17.4. As $\{-\infty, \infty\} \bigcap \mathbb{R} = \emptyset$ and $-\infty \neq \infty$ we have

1. $-\infty \neq \infty$
2. $\forall x \in \mathbb{R}$ we have $-\infty < x$ and $x < \infty$

Theorem 17.5. $\langle \bar{\mathbb{R}}, \leqslant \rangle$ *is fully ordered*

Proof.

reflexitivity. The following cases occurs for $x \in \bar{\mathbb{R}}$

$\boldsymbol{x = \infty}$**.** then by definition $x \leqslant x$

$\boldsymbol{x = -\infty}$**.** then by definition $x \leqslant x$

$\boldsymbol{x} \in \mathbb{R}$. then as $x \leqslant_{\mathbb{R}} x \Rightarrow x \leqslant x$

proving reflexitivity.

anti-symmetry. Let $x, y \in \bar{\mathbb{R}}$ with $x \leqslant y \wedge y \leqslant x$ then the following cases must be considered for $x, y \in \bar{\mathbb{R}}$:

$\boldsymbol{x} = \boldsymbol{y} = \infty$. then $x = y$

$\boldsymbol{x} = \boldsymbol{y} = -\infty$. then $x = y$

$\boldsymbol{x} = \infty \wedge \boldsymbol{y} = -\infty$. then by definition $x \nleqslant y$ so this case does not count

$\boldsymbol{x} = -\infty \wedge \boldsymbol{y} = \infty$. then by definition $y \nleqslant x$ so this case does not count

$\boldsymbol{x} = \infty \wedge \boldsymbol{y} \in \mathbb{R}$. then by definition $x \nleqslant y$ so this case does not count

$\boldsymbol{x} = -\infty \wedge \boldsymbol{y} \in \mathbb{R}$. then by definition $y \nleqslant x$ so this case does not count

$\boldsymbol{x} \in \mathbb{R} \wedge \boldsymbol{y} = \infty$. then by definition $y \nleqslant x$ so this case does not count

$\boldsymbol{x} \in \mathbb{R} \wedge \boldsymbol{y} = -\infty$. then by definition $x \nleqslant y$ so this case does not count

$\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}$. then $x \leqslant_{\mathbb{R}} y \wedge y \leqslant_{\mathbb{R}} x \Rightarrow x = y$

so in all the cases where $x \leqslant y \wedge y \leqslant x$ we have $x = y$

transitivity. Let $x, y, z \in \bar{\mathbb{R}}$ with $x \leqslant y \wedge y \leqslant z$ then the following cases must be considered for $x, y, z \in \bar{\mathbb{R}}$:

$\boldsymbol{x} = \infty \wedge \boldsymbol{y} = \infty \wedge \boldsymbol{z} = \infty$. then $x \leqslant z$

$\boldsymbol{x} = -\infty \wedge \boldsymbol{y} = \infty \wedge \boldsymbol{z} = \infty$. then $x \leqslant z$

$\boldsymbol{x} \in \mathbb{R} \wedge \boldsymbol{y} = \infty \wedge \boldsymbol{z} = \infty$. then $x \leqslant z$

$\boldsymbol{x} = \infty \wedge \boldsymbol{y} = -\infty \wedge \boldsymbol{z} = \infty$. then as $x \nleqslant y$ this cases does not count

$\boldsymbol{x} = -\infty \wedge \boldsymbol{y} = -\infty \wedge \boldsymbol{z} = \infty$. then $x \leqslant z$

$\boldsymbol{x} \in \mathbb{R} \wedge \boldsymbol{y} = -\infty \wedge \boldsymbol{z} = \infty$. then as $x \nleqslant y$ this case does not count

$\boldsymbol{x} = \infty \wedge \boldsymbol{y} \in \mathbb{R} \wedge \boldsymbol{z} = \infty$. then as $x \nleqslant y$ this case does not count

$\boldsymbol{x} = -\infty \wedge \boldsymbol{y} \in \mathbb{R} \wedge \boldsymbol{z} = \infty$. then $x \leqslant z$

$\boldsymbol{x} \in \mathbb{R} \wedge \boldsymbol{y} \in \mathbb{R} \wedge \boldsymbol{z} = \infty$. then $x \leqslant z$

$\boldsymbol{x} = \infty \wedge \boldsymbol{y} = \infty \wedge \boldsymbol{z} = -\infty$. then as $y \nleqslant z$ this case does not count

$\boldsymbol{x} = -\infty \wedge \boldsymbol{y} = \infty \wedge \boldsymbol{z} = -\infty$. then as $y \nleqslant z$ this case does not count

$\boldsymbol{x} \in \mathbb{R} \wedge \boldsymbol{y} = \infty \wedge \boldsymbol{z} = -\infty$. then as $y \nleqslant z$ this case does not count

$\boldsymbol{x} = \infty \wedge \boldsymbol{y} = -\infty \wedge \boldsymbol{z} = -\infty$. then as $x \nleqslant y$ this case does not count

$\boldsymbol{x} = -\infty \wedge \boldsymbol{y} = -\infty \wedge \boldsymbol{z} = -\infty$. then $x \leqslant z$

$\boldsymbol{x} \in \mathbb{R} \wedge \boldsymbol{y} = -\infty \wedge \boldsymbol{z} = -\infty$. then as $x \nleqslant y$ this case does not count

$\boldsymbol{x} = \infty \wedge \boldsymbol{y} \in \mathbb{R} \wedge \boldsymbol{z} = -\infty$. then as $y \nleqslant z$ this case does not count

$\boldsymbol{x} = -\infty \wedge \boldsymbol{y} \in \mathbb{R} \wedge \boldsymbol{z} = -\infty$. then as $y \nleqslant z$ this case does not count

$\boldsymbol{x} \in \mathbb{R} \wedge \boldsymbol{y} \in \mathbb{R} \wedge \boldsymbol{z} = -\infty$. then as $y \nleqslant z$ this case does not count

$\boldsymbol{x=\infty\wedge y=\infty\wedge z\in\mathbb{R}}$. then as $y\nleqslant z$ this case does not count

$\boldsymbol{x=-\infty\wedge y=\infty\wedge z\in\mathbb{R}}$. then as $y\nleqslant z$ this case does not count

$\boldsymbol{x\in\mathbb{R}\wedge y=\infty\wedge z\in\mathbb{R}}$. then as $y\nleqslant z$ this case does not count

$\boldsymbol{x=\infty\wedge y=-\infty\wedge z\in\mathbb{R}}$. then as $x\nleqslant y$ this case does not count

$\boldsymbol{x-\infty\wedge y=-\infty\wedge z\in\mathbb{R}}$. then $x\leqslant z$

$\boldsymbol{x\in\mathbb{R}\wedge y=-\infty\wedge z\in\mathbb{R}}$. then as $x\nleqslant y$ this case does not count

$\boldsymbol{x=\infty\wedge y\in\mathbb{R}\wedge z\in\mathbb{R}}$. then as $x\nleqslant y$ this case does not count

$\boldsymbol{x=-\infty\wedge y\in\mathbb{R}\wedge z\in\mathbb{R}}$. then $x\leqslant z$

$\boldsymbol{x\in\mathbb{R}\wedge y\in\mathbb{R}\wedge z\in\mathbb{R}}$. then $x\leqslant z$

so in all cases that count we have $x\leqslant z$

fully-ordered. The following cases must be considered for $x,y\in\bar{\mathbb{R}}$:

$\boldsymbol{x=y=-\infty}$. $\Rightarrow x\leqslant y$

$\boldsymbol{x=\infty\wedge y=-\infty}$. $\Rightarrow y\leqslant x$

$\boldsymbol{x=-\infty\wedge y=\infty}$. $\Rightarrow x\leqslant y$

$\boldsymbol{x=\infty\wedge y\in\mathbb{R}}$. $\Rightarrow y\leqslant x$

$\boldsymbol{x=-\infty\wedge y\in\mathbb{R}}$. $\Rightarrow x\leqslant y$

$\boldsymbol{x\in\mathbb{R}\wedge y=\infty}$. $\Rightarrow x\leqslant y$

$\boldsymbol{x\in\mathbb{R}\wedge y=-\infty}$. $\Rightarrow y\leqslant x$

$\boldsymbol{x,y\in\mathbb{R}}$. $x\leqslant_{\mathbb{R}}y\vee y\leqslant_{\mathbb{R}}x\Rightarrow x\leqslant y\vee y\leqslant x$ □

Definition 17.6. *$x\in\bar{\mathbb{R}}$ is finite iff $x\in\mathbb{R}$ (or in other words $x\neq\infty,-\infty$)*

We have the following analog to the density of $\mathbb{R}$

Theorem 17.7. *If $x,y\in\bar{\mathbb{R}}$ with $x<y$ then $\exists z\in\bar{\mathbb{R}}$ such that $x<z<y$*

Proof. We have the following possibilities for $x<y\Rightarrow x\neq y$

$\boldsymbol{x\in\mathbb{R}\wedge y=\infty}$. take then $z=x+1\Rightarrow x<z<y$

$\boldsymbol{x=-\infty\wedge y=\infty}$. take then $x=0\Rightarrow x<z<y$

$\boldsymbol{x\in\mathbb{R}\wedge y\in\mathbb{R}}$. using 9.57 there exists then a $z\in\mathbb{R}\subseteq\bar{\mathbb{R}}$ such that $x<x<y$

$\boldsymbol{x=-\infty\wedge y\in\mathbb{R}}$. take then $z=y-1$ so that $x<z<y$ □

Notation 17.8. *Let $\emptyset\neq A\subseteq\mathbb{R}$ be a non empty real set then* $\sup_{\mathbb{R}}(A),\inf_{\mathbb{R}}(A)$ *is the supremum and infinum of A in $\langle\mathbb{R},\leqslant_{\mathbb{R}}\rangle$ (where $\leqslant_{\mathbb{R}}$ is the canonical order in $\mathbb{R}$ see 8.36)*

Notation 17.9. *Let $\emptyset\neq A\subseteq\bar{\mathbb{R}}$ be a non empty extended real set then* $\sup(A)$, $\inf(A)$ *is the supremum and infinum of A in $\langle\bar{\mathbb{R}},\leqslant\rangle$*

We prove now that in $\langle\bar{\mathbb{R}},\leqslant\rangle$ every on empty set has a supremum and infinum

Theorem 17.10. *Let* $\emptyset \neq A \subseteq \bar{\mathbb{R}}$ *then* $\sup(A)$ *and* $\inf(A)$ *exist*

Proof. From the full ordering of $\langle \bar{\mathbb{R}}, \leqslant \rangle$ we have by 2.182 that $\langle \bar{\mathbb{R}}, \leqslant \rangle$ is well ordered and thus by 2.184 conditionally complete. Now if A is non empty then it is bounded above by ∞ and bounded below by $-\infty$ so that by conditional completeness (see 2.175 and the equivalence of inf and sup (see 2.176)) A has a infinum and supremum. □

Theorem 17.11. *Let* $\emptyset \neq A \subseteq \bar{\mathbb{R}}$ *then* $\sup(A) = \infty$ *(or* $\inf(A) = -\infty$*) iff either* $\infty \in A$ *or* $A \bigcap \mathbb{R}$ *is not bounded above (or either* $-\infty \in A$ *or* $A \bigcap \mathbb{R}$ *is not bounded below) in* $\mathbb{R}$

Proof.

⇒. If $\sup(A) = \infty$ then we have for A either

- $\boldsymbol{\infty \in A}$**.** proving $\infty \in A$
- $\boldsymbol{\infty \notin A}$**.** consider then the following cases
  - $\boldsymbol{A = \{-\infty\}}$**.** then as $\sup(\{-\infty\}) = -\infty \neq \infty$ we have a contradiction, so this case can not occur
  - $\boldsymbol{A \neq \{-\infty\}}$**.** then if $A \bigcap \mathbb{R}$ is bounded above in $\mathbb{R}$ we have by 9.43 that $\sup_{\mathbb{R}}(A \bigcap \mathbb{R}) \in \mathbb{R}$ exists. If now $x \in A$ then either $x = -\infty$ and thus $x = -\infty < \sup_{\mathbb{R}}(A \bigcap R)$ or $x \neq -\infty$ and $x \in A \bigcap \mathbb{R} \Rightarrow x \leqslant \sup_{\mathbb{R}}(A \bigcap \mathbb{R})$ so A is bounded above by $\sup_{\mathbb{R}}(A \bigcap \mathbb{R})$ proving that $\sup(A) \leqslant \sup_{\mathbb{R}}(A \bigcap \mathbb{R}) < \infty$ a contradition. So we must have that $A \bigcap \mathbb{R}$ is not bounded above.

If $\inf(A) = -\infty$ then we have for A either

- $\boldsymbol{-\infty \in A}$**.** proving $-\infty$
- $\boldsymbol{-\infty \notin A}$**.** consider then the following cases
  - $\boldsymbol{A = \{\infty\}}$**.** then as $\inf(\{\infty\}) = \infty \neq -\infty$ we have a contradiction, so this case can not occur
  - $\boldsymbol{A \neq \{\infty\}}$**.** then if $A \bigcap \mathbb{R}$ is bounded below in $\mathbb{R}$ we have by 9.43 and 2.176 that $\inf_{\mathbb{R}}(A \bigcap \mathbb{R}) \in \mathbb{R}$ exists. If now $x \in A$ then either $x = \infty \Rightarrow \inf_{\mathbb{R}}(A \bigcap R) < \infty = x$ or $x \neq \infty \Rightarrow x \in A \bigcap \mathbb{R} \Rightarrow \inf_{\mathbb{R}}(A \bigcap \mathbb{R}) \leqslant x$, so A is bounded below by $\inf_{\mathbb{R}}(A \bigcap \mathbb{R})$ proving that $-\infty < \inf_{\mathbb{R}}(A \bigcap \mathbb{R}) \leqslant \inf(A)$ a contradiction. So we must have that $A \bigcap \mathbb{R}$ is not bounded below.

⇐. If $\infty \in A$ then $\infty \leqslant \sup(A) \Rightarrow \sup(A) = \infty$ [if $\sup(A) \neq \infty \Rightarrow \infty \leqslant \sup(A) < \infty$ a contradiction], else if $A \bigcap \mathbb{R}$ is not bounded above in $\mathbb{R}$ then if $\sup(A) < \infty$ we have that $\sup(A \bigcap R) \leqslant \sup(A) < \infty$ giving the contradiction that $A \bigcap \mathbb{R}$ is bounded above, so $\sup(A) = \infty$. If $-\infty \in A$ then $\inf(A) \leqslant -\infty \Rightarrow \inf(A) = -\infty$ [if $\inf(A) \neq -\infty \Rightarrow -\infty < \inf(A) \leqslant -\infty$ a contradiction], else if $A \bigcap \mathbb{R}$ is not bounded below in $\mathbb{R}$ then if $-\infty < \inf(A)$ we have that $-\infty < \inf(A) \leqslant \inf(A \bigcap \mathbb{R})$ giving the contradiction that $A \bigcap \mathbb{R}$ is bounded below, so $\inf(A) = -\infty$. □

We show now that the supremum (infinum) in $\bar{\mathbb{R}}$ is a extension of the supremum and infinum in $\mathbb{R}$.

Theorem 17.12. *Let $\emptyset \neq A \subseteq \bar{\mathbb{R}}$ then*

1. *If $A \subseteq]-\infty, \infty]$ and $\sup(A) < \infty$ then $A \subseteq \mathbb{R}$, $\sup_{\mathbb{R}}(A)$ exists and $\sup(A) = \sup_{\mathbb{R}}(A)$*
2. *If $A \subset -[-\infty, \infty[$ $-\infty < \inf(A)$ then $A \subseteq \mathbb{R}$, $\inf_{\mathbb{R}}(A)$ and $\inf(A) = \inf_{\mathbb{R}}(A)$*

Proof.

1. If $\infty \in A$ then by the above theorem (see 17.11) we have that $\sup(A) = \infty$ a contradiction so we have
$$A \subseteq \mathbb{R}$$
Take $x \in A$ then by definition of the supremum we have $x \leqslant \sup(A) \underset{x \in A \subseteq \mathbb{R}}{\Rightarrow} x \leqslant_{\mathbb{R}} \sup(A)$ proving that A is bounded above by $\sup(A) \in \mathbb{R}$ and thus by conditionally completeness of $\mathbb{R}$ (see 9.43) that $\sup_{\mathbb{R}}(A)$ exists and that $\sup_{\mathbb{R}}(A) \leqslant_{\mathbb{R}} \sup(A)$. If now $\sup_{\mathbb{R}}(A) <_{\mathbb{R}} \sup(A)$ then $\exists a \in A$ such that $\sup(A) < a \leqslant \sup_{\mathbb{R}}(A) \underset{a \in A \subseteq \mathbb{R}}{\Rightarrow} \sup(A) <_{\mathbb{R}} a$ contradicting $a \leqslant_{\mathbb{R}} \sup_{\mathbb{R}}(A)$, hence we have $\sup_{\mathbb{R}}(A) = \sup(A)$
2. If $-\infty \in A$ then by the above theorem (see 17.11) we have that $\sup(A) = \infty$ a contradiction so we have
$$A \subseteq \mathbb{R}$$
Take $x \in A$ then by the definition of the supremum we have $\inf(A) \leqslant x \underset{x \in A \subseteq \mathbb{R}}{\Rightarrow} \inf(A) \leqslant_{\mathbb{R}} x$ proving that A is bounded below by $\inf(A) \in \mathbb{R}$ and thus by conditionally completeness of $\mathbb{R}$ (see 9.43) that $\inf_{\mathbb{R}}(A)$ exists and that $\inf(A) \leqslant_{\mathbb{R}} \inf_{\mathbb{R}}(A)$. If now $\inf(A) < \inf_{\mathbb{R}}(A)$ then there exists a $a \in A$ such that $\inf(A) \leqslant a < \inf_{\mathbb{R}}(A) \underset{a \in A \subseteq \mathbb{R}}{\Rightarrow} a <_{\mathbb{R}} \inf_{\mathbb{R}}(A)$ contradicting $\inf_{\mathbb{R}}(A) \leqslant_{\mathbb{R}} \inf(A)$, hence we have $\inf_{\mathbb{R}}(A) = \inf(A)$. □

Note 17.13. The condition $A \subseteq]-\infty, \infty]$ in (1) above or $A \subseteq [-\infty, \infty[$ in (2) is needed for if $A = \{-\infty\}$ then $A \nsubseteq \mathbb{R}$ and $\sup(A) = -\infty \neq \sup_{\mathbb{R}}(A)$ and if $A = \{\infty\}$ then $A \subseteq \mathbb{R}$ and $\inf(A) = \infty \neq \inf_{\mathbb{R}}(A)$.

Convention 17.14. For the rest of the document, if $A \subseteq \mathbb{R}$ then $\sup(A) < \infty$ means $\sup_{\mathbb{R}}(A)$ exists and $-\infty < \inf(A)$ means $\inf_{\mathbb{R}}(A)$ exists. Where $\sup_{\mathbb{R}}(A) \in \mathbb{R}$ and $\inf_{\mathbb{R}}(A) \in \mathbb{R}$. This is of course consistent with the above theorem.

We extend the algebraic operations on $\mathbb{R}$ as follows (note that this will not make $\langle \bar{\mathbb{R}}, +, \cdot \rangle$ a field)

Definition 17.15. *On $\bar{\mathbb{R}}$ we have the following possibilities for the $+$ and . operations*

1. *Given $x, y \in \bar{\mathbb{R}}$ we have for $x + y$ the following possibilities*

   $x \in \mathbb{R} \wedge y \in \mathbb{R}$. *then $x + y$ is equal to $x + y$ as defined in $\mathbb{R}$*

$\boldsymbol{x \in \mathbb{R} \wedge y = \infty}$. *then* $x + y = \infty$

$\boldsymbol{x \in \mathbb{R} \wedge y = -\infty}$. *then* $x + y = -\infty$

$\boldsymbol{x = \infty \wedge y \in \mathbb{R}}$. *then* $x + y = \infty$

$\boldsymbol{x = \infty \wedge y = \infty}$. *then* $x + y = \infty$

$\boldsymbol{x = \infty \wedge y = -\infty}$. *then* $x + y$ *is* ***not defined***

$\boldsymbol{x = -\infty \wedge y \in \mathbb{R}}$. *then* $x + y = -\infty$

$\boldsymbol{x = -\infty \wedge y = \infty}$. *then* $x + y$ *is* ***not definied***

$\boldsymbol{x = -\infty \wedge y = -\infty}$. *then* $x + y = -\infty$

2. *Given* $x, y \in \bar{\mathbb{R}}$ *we have for* $x \cdot y$ *the following possibilities where* $\mathbb{R}_+ = \{x \in \mathbb{R} | 0 < x\}$, $\mathbb{R}_- = \{x \in \mathbb{R} | x < 0\}$

$\boldsymbol{x \in \mathbb{R} \wedge y \in \mathbb{R}}$. *then* $x \cdot y$ *is equal to* $x \cdot y$ *as defined in* $\mathbb{R}$

$\boldsymbol{x \in \mathbb{R}_+ \wedge y = \infty}$. *then* $x \cdot y = \infty$

$\boldsymbol{x \in \mathbb{R}_- \wedge y = \infty}$. *then* $x \cdot y = -\infty$

$\boldsymbol{x = 0 \wedge y = \infty}$. *then* $x \cdot y = 0$

$\boldsymbol{x = \infty \wedge y = \infty}$. *then* $x \cdot y = \infty$

$\boldsymbol{x = -\infty \wedge y = \infty}$. *then* $x \cdot y = -\infty$

$\boldsymbol{x \in \mathbb{R}_+ \wedge y = -\infty}$. *then* $x \cdot y = -\infty$

$\boldsymbol{x \in \mathbb{R}_- \wedge y = -\infty}$. *then* $x \cdot y = \infty$

$\boldsymbol{x = 0 \wedge y = -\infty}$. *then* $x \cdot y = 0$

$\boldsymbol{x = \infty \wedge y = -\infty}$. *then* $x \cdot y = -\infty$

$\boldsymbol{x = -\infty \wedge y = -\infty}$. *then* $x \cdot y = \infty$

$\boldsymbol{x = \infty \wedge y \in \mathbb{R}_+}$. *then* $x \cdot y = \infty$

$\boldsymbol{x = \infty \wedge y \in \mathbb{R}_-}$. *then* $x \cdot y = -\infty$

$\boldsymbol{x = \infty \wedge y = 0}$. *then* $x \cdot y = 0$

$\boldsymbol{x = \infty \wedge y = \infty}$. *then* $x \cdot y = \infty$

$\boldsymbol{x = \infty \wedge y = -\infty}$. *then* $x \cdot y = -\infty$

$\boldsymbol{x = -\infty \wedge y \in \mathbb{R}_+}$. *then* $x \cdot y = -\infty$

$\boldsymbol{x = -\infty \wedge y \in \mathbb{R}_-}$. *then* $x \cdot y = \infty$

$\boldsymbol{x = -\infty \wedge y = 0}$. *then* $x \cdot y = 0$

$\boldsymbol{x = -\infty \wedge y = \infty}$. *then* $x \cdot y = -\infty$

$\boldsymbol{x = -\infty \wedge y = -\infty}$. *then* $x \cdot y = \infty$

3. *Given* $x \in \bar{\mathbb{R}}$ *we have for* $-x$ *the following possibilities*

$\boldsymbol{x \in \mathbb{R}}$. *then* $-x$ *is as is defined in* $\mathbb{R}$

$\boldsymbol{x = \infty}$. *then* $-x = -\infty$

$\boldsymbol{x = -\infty}$. *then* $-x = \infty$

So in all cases the product of two numbers is defined, however the sum is not defined for the cases $\infty + (-\infty)$ *and* $-\infty + \infty$.

Note 17.16. Using the above definition it is easy to see that $\forall x, y \in \bar{\mathbb{R}}$ we have $x \cdot y = y \cdot x$ and if $x + y$ is defined then $y + x$ is defined and $x + y = y + x$.

Theorem 17.17. *If* $x, y \in \bar{\mathbb{R}}$ *with* $x \leqslant y$ *then* $(-y) \leqslant (-x)$ *and if* $x < y$ *then* $(-y) < (-x)$

Proof. First if $x \leqslant y$ then we have to consider the following possibilities for x

- $\boldsymbol{x \in \mathbb{R}}$**.** then as $x \leqslant y$ we have to consider the following possibilities for y
  - $\boldsymbol{y \in \mathbb{R}}$**.** then using the properties of the reals we have $(-y) \leqslant (-x)$
  - $\boldsymbol{y = \infty}$**.** then $-y = -\infty < -x \in \mathbb{R} \Rightarrow (-y) \leqslant (-x)$
- $\boldsymbol{x = \infty}$**.** then as $x \leqslant y$ we have $y = \infty$ so that $-y = -\infty \leqslant -\infty = -x$ proving $(-y) \leqslant (-x)$
- $\boldsymbol{x = -\infty}$**.** then as $x \leqslant y$ we have to consider the following possibilities
  - $\boldsymbol{y \in \mathbb{R}}$**.** then $-x = \infty$ so that $\mathbb{R} \ni -y \leqslant \infty = -x$ proving $(-y) \leqslant (-x)$
  - $\boldsymbol{y = \infty}$**.** then $-y = -\infty < \infty = -x$ proving $(-y) \leqslant (-x)$
  - $\boldsymbol{y = -\infty}$**.** then $-y = \infty = -x$ proving $(-y) \leqslant (-x)$

this proves the first part of the theorem. For the second part if $x < y$ then $x \neq y$ and $x \leqslant y \Rightarrow (-x) \neq (-y)$ and $(-y) \leqslant (-x)$ proving $(-y) < (-x)$ □

Corollary 17.18. *If* $x, y \in \bar{\mathbb{R}}$ *with* $x \leqslant y$ *then for* $\lambda \in \bar{\mathbb{R}}$ *we have*

1. *If* $0 \leqslant \lambda$ *then* $\lambda \cdot x \leqslant \lambda \cdot y$
2. *If* $\lambda \leqslant 0$ *then* $\lambda \cdot y \leqslant \lambda \cdot x$

Proof.

1. For $0 \leqslant \lambda$ we have the following cases to consider for λ
   - $\boldsymbol{\lambda = 0}$**.** then $\lambda \cdot x = 0 = \lambda \cdot y \Rightarrow \lambda \cdot x \leqslant \lambda \cdot y$
   - $\boldsymbol{\lambda \in \mathbb{R}_+}$**.** then for x we have either
     - $\boldsymbol{x = -\infty}$**.** then $\lambda \cdot x = -\infty \leqslant \lambda \cdot y \Rightarrow \lambda \cdot y \leqslant \lambda \cdot y$
     - $\boldsymbol{x \in \mathbb{R}}$**.** then $\lambda \cdot x \leqslant \lambda \cdot y$ (using 9.41)
     - $\boldsymbol{x = \infty}$**.** then $y \underset{x \leqslant \infty}{=} \infty$ hence $\lambda \cdot x = \infty = \lambda \cdot y \Rightarrow \lambda \cdot x \leqslant \lambda \cdot y$
   - $\boldsymbol{\lambda = \infty}$**.** then for x we have either
     - $\boldsymbol{x = -\infty}$**.** then for y we have either
       - $\boldsymbol{y = -\infty}$**.** $\lambda \cdot x = -\infty = \lambda \cdot y \Rightarrow \lambda \cdot x \leqslant \lambda \cdot y$
       - $\boldsymbol{y \in \mathbb{R}_-}$**.** then $\lambda \cdot x = -\infty = \lambda \cdot y \Rightarrow \lambda \cdot x \leqslant \lambda \cdot y$
       - $\boldsymbol{y = 0}$**.** then $\lambda \cdot x = -\infty \leqslant 0 = \lambda \cdot y \Rightarrow \lambda \cdot x \leqslant \lambda \cdot y$
       - $\boldsymbol{y \in \mathbb{R}_+}$**.** then $\lambda \cdot x = -\infty \leqslant \infty = \lambda \cdot y \Rightarrow \lambda \cdot x \leqslant \lambda \cdot y$

$y=\infty$. then $\lambda\cdot x=-\infty\leqslant\infty=\lambda\cdot y\Rightarrow\lambda\cdot x\leqslant\lambda\cdot y$

$x\in\mathbb{R}_-$. then for y we have either

$y=-\infty$. then $\lambda\cdot x=-\infty=\lambda\cdot y\Rightarrow\lambda\cdot x\leqslant\lambda\cdot y$

$y\in\mathbb{R}_-$. then $\lambda\cdot x=-\infty=\lambda\cdot y\Rightarrow\lambda\cdot x\leqslant\lambda\cdot y$

$y=0$. then $\lambda\cdot x=-\infty\leqslant 0=\lambda\cdot y\Rightarrow\lambda\cdot x\leqslant\lambda\cdot y$

$y\in\mathbb{R}_+$. then $\lambda\cdot x=-\infty\leqslant\infty=\lambda\cdot y\Rightarrow\lambda\cdot x\leqslant\lambda\cdot y$

$y=\infty$. then $\lambda\cdot x=-\infty\leqslant\infty=\lambda\cdot y\Rightarrow\lambda\cdot x\leqslant\lambda\cdot y$

$x=0$. then as $0=x\leqslant y$ we have for y either

$y=0$. then $\lambda\cdot x=0=\lambda\cdot y\Rightarrow\lambda\cdot x\leqslant\lambda\cdot y$

$y\in\mathbb{R}_+$. then $\lambda\cdot x=0\leqslant\infty=\lambda\cdot y\Rightarrow\lambda\cdot x\leqslant\lambda\cdot y$

$y=\infty$. then $\lambda\cdot x=0\leqslant\infty=\lambda\cdot y\Rightarrow\lambda\cdot x\leqslant\lambda\cdot y$

$x\in\mathbb{R}_+$. then as $0<x\leqslant y$ we have either

$y\in\mathbb{R}_+$. then $\lambda\cdot x=\infty=\lambda\cdot y\Rightarrow\lambda\cdot x\leqslant\lambda\cdot y$

$y=\infty$. then $\lambda\cdot x=\infty=\lambda\cdot y\Rightarrow\lambda\cdot x\leqslant\lambda\cdot y$

$x=\infty$. then $y=\infty$ and $\lambda\cdot x=\infty=\lambda\cdot y\Rightarrow\lambda\cdot x\leqslant\lambda\cdot y$

2. As $\lambda\leqslant 0$ we have by 17.17 that $0\leqslant(-\lambda)$ so by (1) we have that $(-\lambda)\cdot x\leqslant(-\lambda)\Rightarrow-(\lambda\cdot x)\leqslant-(\lambda\cdot y)\underset{17.17}{\Rightarrow}\lambda\cdot y\leqslant\lambda\cdot x$ □

We have to be very carefull if we do sums on elements of $\bar{\mathbb{R}}$ for example if $x=10$, $y=4$ and $z=\infty$ then $x+z=y+z=\infty$ however $x\neq y$ also $x+z\leqslant y+z$ however $x\nleqslant y$. The following theorems shows what is possible.

Theorem 17.19. *If $x,y\in\bar{\mathbb{R}}$ and $z\in\mathbb{R}$ then we have*

1. *if $x+z=y+z$ we have $x=y$*
2. *if $x+z\leqslant y+z$ we have $x\leqslant y$*
3. *if $x+z<y+z$ we have $x<y$*
4. *if $x+z=y$ we have $x=y-z$*
5. *if $x+z\leqslant y$ we have $x\leqslant y+(-z)$*
6. *if $x+z<y$ we have $x<y+(-z)$*
7. *if $x\leqslant y+z$ we have $x+(-z)\leqslant y$*
8. *if $x<y+z$ we have $x+(-z)<y$*

Proof.

1. Assume $x+z=y+z$ and consider the following possibilities for x,y when

$x\in\mathbb{R}\wedge y\in\mathbb{R}$. then as $x+z=y+z$ we have $(x+z)-z=(y+z)-z\Rightarrow x=y$

$x\in\mathbb{R}\wedge y=\infty$. then $x+z\neq y+z=\infty$ so this does not apply

$x\in\mathbb{R}\wedge y=-\infty$. then $x+z\neq y+z=-\infty$ so this does not apply

$x = \infty \wedge y \in \mathbb{R}$. then $\infty = x + z \neq y + z$ so this does not apply

$x = \infty \wedge y = \infty$. then $x = y$

$x = \infty \wedge y = -\infty$. then $\infty = x + z \neq y + z = -\infty$ so this does not apply

$x = -\infty \wedge y \in \mathbb{R}$. then $-\infty = x + z \neq y + z$ so this does not apply

$x = -\infty \wedge y = \infty$. then $-\infty = x + z \neq y + z = \infty$ so this does not apply

$x = -\infty \wedge y = -\infty$. then $x = y$

so in all cases that apply we have proved that $x = y$

2. Assume $x + y \leqslant y + z$ and consider the following possibilities for x, y

   $x \in \mathbb{R} \wedge y \in \mathbb{R}$. then using the properties of the reals we have $x \leqslant y$

   $x \in \mathbb{R} \wedge y = \infty$. then $x \leqslant y$

   $x \in \mathbb{R} \wedge y = -\infty$. then $x + z \nleqslant y + z = -\infty$ so this does not apply

   $x = \infty \wedge y \in \mathbb{R}$. then $\infty = x + z \nleqslant y + z$ so this does not apply

   $x = \infty \wedge y = \infty$. then $x \leqslant y$

   $x = \infty \wedge y = -\infty$. then $\infty = x + z \nleqslant y + z = -\infty$ so this does not apply

   $x = -\infty \wedge y \in \mathbb{R}$. then $x < y \Rightarrow x \leqslant y$

   $x = -\infty \wedge y = \infty$. then $x < y \Rightarrow x \leqslant y$

   $x = -\infty \wedge y = -\infty$. then $x = y \Rightarrow x \leqslant y$

   so in all cases that apply we have $x \leqslant y$

3. Assume that $x + z < y + z$

   $x \in \mathbb{R} \wedge y \in \mathbb{R}$. then as $x + z < y + z$ we have $(x + z) - z < (y + z) - z \Rightarrow x < y$

   $x \in \mathbb{R} \wedge y = \infty$. then $x < y$

   $x \in \mathbb{R} \wedge y = -\infty$. then $x + z \nless y + z = -\infty$ so this does not apply

   $x = \infty \wedge y \in \mathbb{R}$. then $\infty = x + z \nless y + z$ so this does not apply

   $x = \infty \wedge y = \infty$. then $\infty = x + z \nless y + z = \infty$ so this does not apply

   $x = \infty \wedge y = -\infty$. then $\infty = x + z \nless y + z = -\infty$ so this does not apply

   $x = -\infty \wedge y \in \mathbb{R}$. then $x < y$

   $x = -\infty \wedge y = \infty$. then $x < y$

   $x = -\infty \wedge y = -\infty$. then $-\infty = x + z \nless y + z = -\infty$ so this does not apply

   So in all cases that apply we have $x < y$

4. If $x + z = y$ we have for x the following possibilities

   $x \in \mathbb{R}$. then $y \in \mathbb{R}$ and we have $x = y + (-z)$

   $x = \infty$. then $y = \infty$ and we have $x = \infty = \infty + (-z) = y + (-z)$

   $x = -\infty$. then $y = -\infty$ and we have $x = -\infty = -\infty + (-z) = y + (-z)$

so in all cases we have $x = y + (-z)$

5. If $x + z \leqslant y$ we have for x the following possibilities

   $x \in \mathbb{R}$. then for y we have either

   $y \in \mathbb{R}$. then $x \leqslant y + (-z)$

   $y = \infty$. then $x \leqslant \infty = \infty + (-z) \Rightarrow x \leqslant y + (-z)$

   $x = \infty$. then $x + z = \infty \Rightarrow y = \infty$ so that $x = \infty = \infty + (-z) = y + (-z) \Rightarrow x \leqslant y + (-z)$

   $x = -\infty$. then $x + (-z) = -\infty$ and we have the following possibilites for y

   $y \in \mathbb{R}$. then $x = -\infty < y \Rightarrow x \leqslant y + (-z)$

   $y = \infty$. then $x = -\infty < \infty = y + (-z) \Rightarrow x \leqslant y + (-z)$

   $y = -\infty$. then $x = -\infty = -\infty + (-z) = y + (-z) \Rightarrow x \leqslant y + (-z)$

6. If $x + z < y$ we have for x the following possibilities

   $x \in \mathbb{R}$. then we have for either

   $y \in \mathbb{R}$. then by the properties of the reals we have $x < y + (-z)$

   $y = \infty$. then $x < \infty = \infty + (-z) = y + (-z) \Rightarrow x < y + (-z)$

   $x = \infty$. then $x + z = \infty < y$ a contradiction so this does not apply

   $x = -\infty$. then we have the following possibilities for y

   $y \in \mathbb{R}$. then $x = -\infty < y - z \Rightarrow x < y + (-z)$

   $y = \infty$. then $x = -\infty < \infty = \infty + (-z) = y + (-z) \Rightarrow x < y + (-z)$

   so in all cases we have $x < y + (-z)$

7. If $x \leqslant y + z$ then for x we have the following possibilities for x

   $x \in \mathbb{R}$. then for y we have the following possibilities (as $-\infty + z = -\infty < x$)

   $y \in \mathbb{R}$. then $x + (-z) \leqslant y$

   $y = \infty$. then $x + (-z) \leqslant \infty = y \Rightarrow x + (-z) \leqslant y$

   $x = \infty$. then $y + z = \infty$ so that $y = \infty$ and $x + (-z) = \infty = y \Rightarrow x + (-z) \leqslant y$

   $x = -\infty$. then $x + (-z) = -\infty \leqslant y$ proving that $x + (-z) \leqslant y$

   so in all cases we have $x + (-z) \leqslant y$

8. If $x < y + z$ then for x we have the following possibilities for x (as $\infty \nless y + z$)

   $x \in \mathbb{R}$. then $x + (-z) < y$

   $x = -\infty$. then $x + (-z) = -\infty < y \Rightarrow x + (-z) < y$

   so in all cases we have $x + (-z) < y$ □

Theorem 17.20. *If $x, y, z \in \bar{\mathbb{R}}$ such that $x \leqslant y$ then $x + z \leqslant y + z$ assuming that the sums are well defined. If $x < y$ and $z \in \mathbb{R}$ then $x + z < y + z$.*

Proof. As $x+z$, $y+z$ must be well defined we have only the following cases to consider (the others are not well defined)

- $\boldsymbol{x\in\mathbb{R}\wedge y\in\mathbb{R}\wedge z\in\mathbb{R}}$**.** then using the properties of the real numbers we have that $x\leqslant y=x+z\leqslant y+z$
- $\boldsymbol{x\in\mathbb{R}\wedge y\in\mathbb{R}\wedge z=\infty}$**.** then $x+z=\infty=y+z\Rightarrow x+z\leqslant y+z$
- $\boldsymbol{x\in\mathbb{R}\wedge y\in\mathbb{R}\wedge z=-\infty}$**.** then $x+z=-\infty=y+z\Rightarrow x+z\leqslant y+z$
- $\boldsymbol{x\in\mathbb{R}\wedge y=\infty\wedge z\in\mathbb{R}}$**.** then $x+z<\infty=y+z\Rightarrow x+z\leqslant y+z$
- $\boldsymbol{x\in\mathbb{R}\wedge y=\infty\wedge z=\infty}$**.** then $x+z<\infty=y+z\Rightarrow x+z<y+z$
- $\boldsymbol{x\in\mathbb{R}\wedge y=-\infty\wedge x\in\mathbb{R}}$**.** then $x\nleqslant y$ so this does not count
- $\boldsymbol{x\in\mathbb{R}\wedge y=-\infty\wedge x=-\infty}$**.** then $x\nleqslant y$ so this does not count
- $\boldsymbol{x=\infty\wedge y\in\mathbb{R}\wedge z\in\mathbb{R}}$**.** then $x\nleqslant y$ so this does not count
- $\boldsymbol{x=\infty\wedge y\in\mathbb{R}\wedge z=\infty}$**.** then $x+z=\infty=y+z\Rightarrow x+z\leqslant y+z$
- $\boldsymbol{x=\infty\wedge y=\infty\wedge z\in\mathbb{R}}$**.** then $x+z=y+z\Rightarrow x+z\leqslant y+z$
- $\boldsymbol{x=\infty\wedge y=\infty\wedge z=\infty}$**.** then $x+z\leqslant y+z\Rightarrow x+z\leqslant y+z$
- $\boldsymbol{x=\infty\wedge y=-\infty\wedge z\in\mathbb{R}}$**.** then $x\nleqslant y$ this does not count
- $\boldsymbol{x=-\infty\wedge y\in\mathbb{R}\wedge z=-\infty}$**.** then $x+z=-\infty=y+z\Rightarrow x+z\leqslant y+z$
- $\boldsymbol{x=-\infty\wedge y=\infty\wedge z\in\mathbb{R}}$**.** then $x+z=-\infty<\infty=y+z\Rightarrow x+z\leqslant y+z$
- $\boldsymbol{x=-\infty\wedge y=-\infty\wedge z=-\infty}$**.** then $z+z=-\infty=y+z\Rightarrow x+z\leqslant y+z$

proving the first part of the theorem.

For the second part if $x<y\Rightarrow x\leqslant y\underset{\text{first part}}{\Rightarrow}x+z\leqslant y+z$. Assume now that $x+z=y+z$ then as $z\in\mathbb{R}$ we have by the previous theorem (see 17.19) that $x=y$ contradicting $x<y$ so we must have that $x+z<y+z$. □

Theorem 17.21. *Let $a,b\in\bar{\mathbb{R}}$ such that $a,b\geqslant 0$ then we have*

1. *$a+(-b)$ is well defined if and only if $a<\infty\vee b<\infty$*
2. *$a+(-b)=\infty$ if and only if $a=\infty$ and $b<\infty$*
3. *$a+(-b)=-\infty$ if and only $a<\infty\wedge b=\infty$*
4. *$a+(-b)<\infty$ if and only $a<\infty\wedge b<\infty$*
5. *If $a+(-b)$ and $c+(-d)$ are well defined with $a+(-b)=c+(-d)$ then $a+d$, $c+b$ are well defined and $a+d=c+b$*

Proof. We have the following possibilities to consider for a,b where $0\leqslant a,b$

- $\boldsymbol{a<\infty\wedge b<\infty}$**.** then $a,b\in\mathbb{R}$ so that $a+(-b)$ is well defined and $a+(-b)\in\mathbb{R}\Rightarrow a+(-b)<\infty$
- $\boldsymbol{a<\infty\wedge b=\infty}$**.** then $a+(-b)=a+(-\infty)\underset{a\in\mathbb{R}}{=}-\infty$ (by definition)
- $\boldsymbol{a=\infty\wedge b<\infty}$**.** then $a+(-b)=\infty+(-b)\underset{b\in\mathbb{R}=-b\in\mathbb{R}}{=}\infty$ (by definition)
- $\boldsymbol{a=\infty\wedge b=\infty}$**.** then $a+(-b)=\infty+(-\infty)$ which is not well defined

So the only cases where $a+(-b)$ are well defined are these where either $a<\infty$ or $b<\infty$ giving (1). Further $a+(-b)=\infty$ only if $a=\infty\wedge b<\infty$, $a+(-b)=-\infty$ only if $a<\infty\wedge b=\infty$ and $a+(-b)<\infty$ only if $a<\infty\wedge b<\infty$ giving (2),(3) and (4). Finally to prove (5) if $a+(-b), c+(-d)$ are well defined and $a+(-b)=c+(-d)$ then we have by (1) only the following cases to consider for a, b

$a<\infty\wedge b<\infty$. then $a+(-b)<\infty$ and for c, d we have either

- **$c<\infty\wedge d<\infty$.** then $c+(-d)<\infty$, also as $a, b, c, d<\infty$ we have that $a+d, c+b$ are defined. Also from $a+(-b)=c+(-d)$ we have using the associativitiy and commutativity of $+$ in $\mathbb{R}$ that $a+d=c+b$
- **$c<\infty\wedge d=\infty$.** then $c+(-d)=-\infty\neq a+(-b)$ so this case does not apply.
- **$c=\infty\wedge d<\infty$.** then $c+(-d)=\infty\neq a+(-b)$ so this case does not apply

$a<\infty\wedge b=\infty$. then $a+(-b)=-\infty$ and for c, d we have either

- **$c<\infty\wedge d<\infty$.** then $c+(-d)<\infty\Rightarrow c+(-d)\neq a+(-b)$ so this case does not apply
- **$c<\infty\wedge d=\infty$.** then $c+(-d)=-\infty=(a+(-b))$ then $a+d, c+b$ is well defined and $a+d=\infty=c+b$
- **$c=\infty\wedge d<\infty$.** then $c+(-d)=\infty\neq a+(-b)$ so this case does not apply

$a=\infty\wedge b<\infty$. then $a-b=\infty$ and for c, d we have either

- **$c<\infty\wedge d<\infty$.** then $c+(-d)<\infty\Rightarrow c+(-d)\neq a-b$ so this case does not apply
- **$c<\infty\wedge d=\infty$.** then $c+(-d)=-\infty\neq a-b$ so this case does not apply
- **$c=\infty\wedge d<\infty$.** then $c+(-d)=\infty=a-b$ then $a+d, c+b$ is well defined and $a+d=\infty=c+b$

So in all the cases that apply we have $a+d, b+c$ is well defined and $a+d=c+b$ □

Theorem 17.22. *Let $x, y, z\in\bar{\mathbb{R}}$ with $0\leqslant y$, $z=x+y$ (where $x+y$ is well defined) then we have $x\leqslant z$*

Proof. As $x+y$ is well defined we have to consider the following cases (taking in account that $0\leqslant y$)

$x, y\in\mathbb{R}$. as $0\leqslant y$ we have $x=x+0\leqslant x+y$

$x=\infty, y\in\mathbb{R}$. as $z=x+y=\infty=x$ we have $x\leqslant z$

$x=-\infty, y\in\mathbb{R}$. then $x\leqslant z$

$x\in\mathbb{R}, y=\infty$. then $z=x+\infty=\infty$ so that $x\leqslant z$

$x=\infty, y=\infty$. then $z=\infty+\infty=\infty$ so that $x\leqslant z$ □

Note 17.23. To simplify notations we adopt the notation $a-b$ to mean $a+(-b)$ so $a-\infty=a+(-\infty)$

Another combination that we need is the following

Theorem 17.24. *Let $a, b, c, d \in \bar{\mathbb{R}}$ with $a, b, c, d \geqslant 0$ be such that $(a-b)+(c-d)$ is well defined then $(a+c)-(b+d)$ is well defined and $(a-b)+(c-d)=(a+c)-(b+d)$*

Proof. First assume that $(a-b)+(c-d)$ is well defined, then we must consider the following cases for $(a-b)$

$(a-b) \in \mathbb{R}$. then $a-b<\infty$ and thus by the previous theorem $a \leqslant a, b<\infty$ giving $a, b \in \mathbb{R}$ and thus $a-b \in \mathbb{R}$ for $(c-d)$ we have then the following cases to consider

- **$(c-d) \in \mathbb{R}$.** thus $c-d<\infty$ and thus by the previous theorem we have $0 \leqslant c<\infty \wedge 0 \leqslant d<\infty$ so that $c, d \in \mathbb{R}$. As $a, b, c, d \in \mathbb{R}$ we have that $(a+c)-(b+d)$ is well defined and $(a-b)+(c-d)=(a+c)-(b+d)$

- **$(c-d)=\infty$.** then we have $c=\infty$ and $0 \leqslant d<\infty \Rightarrow d \in \mathbb{R}$ (previous theorem). So $(a+c)=\infty$, $(b+d) \in \mathbb{R}$ and thus $(a+c)-(b+d)$ is well defined and $(a+c)-(b+d)=\infty$, as $(a-b)+(c-d)=\infty$ we have also $(a-b)+(c-d)=(a+c)-(b+d)$

- **$(c-d)=-\infty$.** then $0 \leqslant c<\infty$ and $d=\infty$ (previous theorem). So $(a+c) \in \mathbb{R}$, $(b+d)=\infty$ giving the well defined $(a+c)-(b+d)=-\infty=(a-b)+(c-d)$

$(a-b)=\infty$. then we must have $-\infty<(c+d)$, $a=\infty$ and $0 \leqslant b<\infty \Rightarrow b \in \mathbb{R}$ giving the possible cases for $(c-d)$

- **$(c-d) \in \mathbb{R}$.** then $c, d \in \mathbb{R}$ and thus $(a+c)=\infty$ and $(b+d) \in \mathbb{R}$ giving the well defined $(a+c)-(b+d)=\infty=(a-b)+(c-d)$

- **$(c-d)=\infty$.** then $c=\infty$, $0 \leqslant d<\infty \Rightarrow d \in \mathbb{R}$ and thus $(a+c)=\infty$, $(b+d) \in \mathbb{R}$ giving the well defined $(a+c)-(b+d)=\infty=(a-b)+(c-d)$

$(a-b)=-\infty$. then we must have $0 \leqslant a<\infty \Rightarrow a \in \mathbb{R}$, $b=\infty$ and $(c-d)<\infty$ giving the following cases for $(c-d)$

- **$(c-d) \in \mathbb{R}$.** then $c, d \in \mathbb{R}$ so that $a+c \in \mathbb{R}$ and $(b+d)=\infty$ giving the well defined $(a+c)-(b+d)=-\infty=(a-b)+(c-d)$

- **$(c-d)=-\infty$.** then $0 \leqslant c<\infty \Rightarrow c \in \mathbb{R}, d=\infty$ giving $(a+c) \in \mathbb{R}$, $(b+d)=\infty$ and thus the well defined $(a+c)-(b+d)=-\infty=(a-b)+(c-d)$ □

Theorem 17.25. *Let $a, b, c, d \in \bar{\mathbb{R}}$ with $0 \leqslant a, b, c, d$ then if $a-b, c-d$ are well defined with $a-b=c-d$ then $a+d, c+b$ are well defined and $a+d=c+b$*

Proof. We have the following cases to consider for $a-b$

$a-b \in \mathbb{R}$. then from $a-b=c-d$ we have that $c-d \in \mathbb{R}$ hence $a, b, c, d \in \mathbb{R}$ so using the properties of $\mathbb{R}$ we find $a+d=c+b$

$a-b=\infty$. then from $a-b=c-d$ we have that $c-d=\infty$. Using 17.21 we must then have $a=\infty, c=\infty, b, d \in \mathbb{R}$ hence $a+d, c+b$ are well defined and $a+d=\infty=c+b$

$a - b = -\infty$. then from $a - b = c - d$ we have that $c - d = -\infty$. Using 17.21 we must then have $a, c \in \mathbb{R}, b = \infty = d$ hence $a + d, c + b$ are well defined and $a + d = \infty = c + b$ □

Although $\langle \bar{\mathbb{R}}, + \rangle$ is not a group we can prove that $\langle \bar{\mathbb{R}}_+, + \rangle$ where $\bar{\mathbb{R}}_+ = \{x \in \bar{\mathbb{R}} | x \geqslant 0\}$ is a abelian semi-group (see semi-group) which is import because then all the theorems about generalized sums and semi-groups are valid.

Theorem 17.26. *$\langle \bar{\mathbb{R}}_+, + \rangle$ where $\bar{\mathbb{R}}_+ = \{x \in \bar{\mathbb{R}} | x \geqslant 0\}$ forms a abelian semi-group*

Proof. First if $x, y \in \bar{\mathbb{R}}_+$ then we have [as $-\infty < 0$] either

$x = \infty \vee y = \infty$. then $x + y = \infty \in \bar{\mathbb{R}}_+$

$x, y \in \mathbb{R}_+$. then as $0 \leqslant x, y$ we have that $0 \leqslant x \leqslant x + y \Rightarrow x + y \in \bar{\mathbb{R}}_+$

so that $+: \bar{\mathbb{R}}_+ \times \bar{\mathbb{R}}_+ \rightarrow \bar{\mathbb{R}}_+$ defined by $(x, y) \rightarrow x + y$ is indeed a function.

Neutral Element. First $0 \in \bar{\mathbb{R}}_+$, next let $x \in \bar{\mathbb{R}}_+$ then we have [as $-\infty < 0$] either

$x \in \mathbb{R}_+$. then $x + 0 = 0 + x = 0$ [as $\langle \mathbb{R}, + \rangle$ is a group with neutral element 0

$x = \infty$. then $\infty + 0 = \infty = \infty + 0$

so $0 \in \bar{\mathbb{R}}$ is a neutral element.

Associativity. Let $x, y, z \in \bar{\mathbb{R}}$ then we have [as $-\infty < 0$] either

$\exists z \in \{x, y, z\} \vdash z = \infty$.

$x, y, z \in \mathbb{R}_+$. then $x + (y + z) = (x + y) + z$ [as $\langle \mathbb{R}, + \rangle$ is a group

Commutativity. Let $x, y \in \bar{\mathbb{R}}_+$ then as $[-\infty < 0]$ we have either

$x = \infty \vee y = \infty$. and then $x + y = \infty = y + x$

$x, y \in \mathbb{R}_+$. then $x + y = y + x$ [as $\langle \mathbb{R}, + \rangle$ is a Abelian group] □

Theorem 17.27. *Let $x, y, x', y' \in \bar{\mathbb{R}}$ be such that $0 \leqslant x \leqslant x'$ and $0 \leqslant y \leqslant y$ then $x + y \leqslant x' + y'$*

Proof. We have the following possibilities for x', y'

$x' = \infty \vee y' = \infty$. then $x' + y' = \infty$ so that $x + y \leqslant \infty = x' + y'$

$x', y' < \infty$. then $x < \infty, y < \infty$ so that using 9.41 $x \leqslant x' \Rightarrow x + y \leqslant x' + y \wedge y \leqslant y' \Rightarrow x' + y \leqslant x' + y' \Rightarrow x + y \leqslant x' + y'$ □

We have also the distributive law in $\bar{\mathbb{R}}$ as is proved in the following theorem, but we have to be carefull because for example $\infty \cdot (1 - 3) = \infty \cdot (-2) = -\infty$ but $\infty \cdot 1 + \infty \cdot (-3) = \infty + (-\infty)$ which not defined. So we must restrict the distributive law.

Theorem 17.28. *Let $\alpha \in \mathbb{R}$, $x, y \in \bar{\mathbb{R}}$ then if $x + y$ is defined we have $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$*

Proof. $x+y$ is defined only in the following cases

$x, y \in \mathbb{R}$. then $\alpha \cdot (x+y) = \alpha \cdot x + \alpha \cdot y$ by the properties of the real numbers.

$x \in \mathbb{R}, y = -\infty$. then for α we have either

- **$\alpha = 0$.** $\alpha \cdot (x+y) = 0 = 0 + 0 = 0 \cdot x + 0 \cdot y = \alpha \cdot x + \alpha \cdot y$
- **$\alpha < 0$.** $\alpha \cdot (x+y) = \alpha \cdot (-\infty) = \infty = \alpha \cdot x + \infty = \alpha \cdot x + \alpha \cdot (-\infty) = \alpha \cdot x + \alpha \cdot y$
- **$\alpha > 0$.** $\alpha \cdot (x+y) = \alpha \cdot \infty = -\infty = \alpha \cdot x + (-\infty) = \alpha \cdot x + \alpha \cdot (-\infty) = \alpha \cdot x + \alpha \cdot y$

$x \in \mathbb{R}, y = \infty$. then for α we have either

- **$\alpha = 0$.** $\alpha \cdot (x+y) = 0 = 0 + 0 = 0 \cdot x + 0 \cdot y = \alpha \cdot x + \alpha \cdot y$
- **$\alpha < 0$.** $\alpha \cdot (x+y) = \alpha \cdot \infty = -\infty = \alpha \cdot x + (-\infty) = \alpha \cdot x + \alpha \cdot \infty = \alpha \cdot x + \alpha \cdot y$
- **$\alpha > 0$.** $\alpha \cdot (x+y) = \alpha \cdot \infty = \infty = \alpha \cdot x + \infty = \alpha \cdot x + \alpha \cdot \infty = \alpha \cdot x + \alpha \cdot x + \alpha \cdot y$

$x = \infty, y \in \mathbb{R}$. then for α we have either

- **$\alpha = 0$.** $\alpha \cdot (x+y) = 0 = 0 + 0 = 0 \cdot x + 0 \cdot y = \alpha \cdot x + \alpha \cdot y$
- **$\alpha < 0$.** $\alpha \cdot (x+y) = \alpha \cdot \infty = -\infty = -\infty + \alpha \cdot y = \alpha \cdot \infty + \alpha \cdot y = \alpha \cdot x + \alpha \cdot y$
- **$\alpha > 0$.** $\alpha \cdot (x+y) = \alpha \cdot \infty = \infty = \infty + \alpha \cdot y = \alpha \cdot \infty + \alpha \cdot y = \alpha \cdot x + \alpha \cdot y$

$x = -\infty, y \in \mathbb{R}$. then for α we have either

- **$\alpha = 0$.** $\alpha \cdot (x+y) = 0 = 0 + 0 = 0 \cdot x + 0 \cdot y = \alpha \cdot x + \alpha \cdot y$
- **$\alpha < 0$.** $\alpha \cdot (x+y) = \alpha \cdot (-\infty) = \infty = \infty + \alpha \cdot y = \alpha \cdot (-\infty) + \alpha \cdot y = \alpha \cdot x + \alpha \cdot y$
- **$\alpha > 0$.** $\alpha \cdot (x+y) = \alpha \cdot (-\infty) = -\infty = -\infty + \alpha \cdot y = \alpha \cdot (-\infty) + \alpha \cdot y = \alpha \cdot x + \alpha \cdot y$

$x = \infty, y = \infty$. then for α we have either

- **$\alpha = 0$.** $\alpha \cdot (x+y) = 0 = 0 + 0 = 0 \cdot x + 0 \cdot y = \alpha \cdot x + \alpha \cdot y$
- **$\alpha < 0$.** $\alpha \cdot (x+y) = \alpha \cdot \infty = -\infty = -\infty + (-\infty) = \alpha \cdot \infty + \alpha \cdot \infty = \alpha \cdot x + \alpha \cdot y$
- **$\alpha > 0$.** $\alpha \cdot (x+y) = \alpha \cdot \infty = \infty = \infty + \infty = \alpha \cdot \infty + \alpha \cdot \infty = \alpha \cdot x + \alpha \cdot y$

$x = -\infty, y = -\infty$. then for α we have either

- **$\alpha = 0$.** $\alpha \cdot (x+y) = 0 = 0 + 0 = 0 \cdot x + 0 \cdot y = \alpha \cdot x + \alpha \cdot y$
- **$\alpha < 0$.** $\alpha \cdot (x+y) = \alpha \cdot (-\infty) = \infty = \infty + \infty = \alpha \cdot (-\infty) + \alpha \cdot (-\infty) = \alpha \cdot x + \alpha \cdot y$
- **$\alpha > 0$.** $\alpha \cdot (x+y) = \alpha \cdot \infty = \infty = \infty + \infty = \alpha \cdot \infty + \alpha \cdot \infty = \alpha \cdot x + \alpha \cdot y$

So in all cases we have $\alpha \cdot (x+y) = \alpha \cdot x + \alpha \cdot y$ □

Theorem 17.29. *Let $\emptyset \neq A \subseteq \bar{\mathbb{R}}$ then we have*

1. *If $y \in \mathbb{R}$ then*

   a. $\sup(A) + y$ *is well defined and* $\sup(\{x+y | x \in A\}) = \sup(A) + y$

b. $\inf(A)+y$ *is well defined and* $\inf(\{x+y|x\in A\})=\inf(A)+y$

2. *If* $y\in[0,\infty]$ *and* $A\subseteq[0,\infty]$ *then* $\sup(A)+y$ *is well defined and* $\sup(A+y)=\sup(A)+y$

3. *If* $y\in[-\infty,0]$ *and* $A\subseteq[-\infty,0]$ *then* $\inf(A)+y$ *is well defined and* $\inf(A+y)=\inf(A)+y$

$y\in\mathbb{R}$ *then we have that* $\sup(\{x+y|x\in A\})=\sup(A)+y$ *and* $\inf(\{x+y|x\in A\})=\inf(A)+y$ *where the last sums are well defined.*

Proof. As $y\in\mathbb{R}$ we have that $\forall x\in\mathbb{R}$ $x+y$ is well defined and $\sup(A)+y$, $\inf(A)+y$ is well defined.

1.

a. Now $\forall x\in A$ we have $x\leqslant\sup(A)$ so that by 17.20 we have $x+y\leqslant\sup(A)+y$ so that

$$\sup(\{x+y|x\in A\})\leqslant\sup(A)+y \tag{17.1}$$

Assume now that $\sup(\{x+y|y\in A\})<\sup(A)+y\underset{y\in\mathbb{R}\wedge 17.20}{\Rightarrow}\sup(\{x+y|y\in A\})-y<\sup(A)$ so by the definition of the supremum there exists a $z\in A$ such that $\sup(\{x+y|x\in A\})-y<z\leqslant\sup(A)\underset{y\in\mathbb{R}\text{and}17.19}{\Rightarrow}\sup(\{x+y|x\in A\})<z+y$ a contradiction as $z+y\in\{x+y|y\in A\}$. So we must have that $\sup(A)+y\leqslant\sup(\{x+y|x\in A\})$ which together with 17.1 gives

$$\sup(\{x+y|x\in A\})=\sup(A)+y$$

b. Now $\forall x\in A$ we have $\inf(A)\leqslant x$ so that by 17.20 we have $\inf(A)+y\leqslant x+y$ proving that

$$\inf(A)+y\leqslant\inf(\{x+y|x\in A\}) \tag{17.2}$$

Assume now that $\inf(A)+y<\inf(\{x+y|x\in A|\})$ then using 17.19 we have $\inf(A)<\inf(\{x+y|x\in A\})-y$ so by the definition of the infinum there exists a $z\in A$ such that $\inf(A)\leqslant z<\inf(\{x+y|x\in A|\})-y\underset{y\in\mathbb{R}\text{and}17.19}{\Rightarrow}z+y<\inf(\{x+y|x\in A\})$ a contradiction as $z+y\in\{x+y|x\in A\}$. So we must have that $\text{in}(\{x+y|x\in A\})\leqslant\inf(A)+y$ proving that

$$\inf(A)+y=\inf(\{x+y|x\in A|\})$$

2. For $y\in[0,\infty]$ we have either

$y\in\mathbb{R}$. then by 1.a we have that $\sup(A)+y$ is well defined and $\sup(\{x+y|x\in A\})=\sup(A)+y$

$y=\infty$. then as $A\subseteq[0,\infty]$ we have that $\forall x\in A+y$ that $x=a+y=a+\infty=\infty$ so that $\sup(A+y)=\infty$, further $0\leqslant\sup(A)\Rightarrow\sup(A)+y=\sup(A)+\infty$ is well defined and $\sup(A+y)=\infty=\sup(A)+y$

3. For $y\in[-\infty,0]$ we have either

$y\in\mathbb{R}$. then by 1.b we have that $\inf(A)+y$ is well defined and $\inf(\{x+y|x\in A\})=\inf(A)+y$

$y=\infty$. then as $A\subseteq[-\infty,0]$ we have that $\forall x\in A+y$ that $x=a+y=a+(-\infty)=-\infty$ so that $\inf(A+y)=-\infty$, further $\inf(A)\leqslant 0\Rightarrow\inf(A)+y=\inf(A)+(-\infty)$ is well defined and $\inf(A+y)=-\infty=\inf(A)+y$ □

Theorem 17.30. *Let $\emptyset\neq A\subseteq\bar{\mathbb{R}}$ be a non empty set then*

1. $-\sup(A)=\inf(\{-x|x\in A\})$
2. $-\inf(A)=\sup(\{-x|x\in A\})$
3. *If $\alpha\in[0,\infty[$ then $\alpha\cdot\sup(A)=\sup(\{\alpha\cdot x|x\in A\})$*
4. *If $\alpha\in[0,\infty[$ then $\alpha\cdot\inf(A)=\inf(\{\alpha\cdot x|x\in A\})$*

Proof.

1. If $x\in A$ then $x\leqslant\sup(A)\underset{17.17}{\Rightarrow}-\sup(A)\leqslant -x$ so that

$$-\sup(A)\leqslant\inf(\{-x|x\in A\}) \tag{17.3}$$

Assume now that $-\sup(A)<\inf(\{-x|x\in A\})$ then using 17.17 we have $-\inf(\{-x|x\in A\})<\sup(A)$ so by the definition of the supremum there exists a $z\in A$ such that $-\inf(\{-x|x\in A\})<z\leqslant\sup(A)\underset{17.17}{\Rightarrow}-z<\inf(\{-x|x\in A\})$ which as $-z\in\{-x|x\in A\}$ is a contradiction. So we must have $\inf(\{-x|x\in A\})\leqslant-\sup(A)$ which together with 17.3 proves that

$$-\sup(A)=\inf(\{-x|x\in A\})$$

2. Take $B=\{-x|x\in A\}$ then $A=\{-x|x\in B\}$ and $-\inf(A)=-\inf(\{-x|x\in B\})\underset{(1)}{=}-(-\sup(B))=\sup(\{-x|x\in A\})$ proving

$$-\inf(A)=\sup(\{-x|x\in A\})$$

3. We must consider two cases for α

$\alpha=0$. In this cases $\{\alpha\cdot x|x\in A\}=\{0\}$ and thus $\sup(\{\alpha\cdot x|x\in A\})=\sup(\{0\})=0=0\cdot\sup(A)=\alpha\cdot\sup(A)$

$0\neq\alpha$. $\forall x\in A$ we have have $x\leqslant\sup(A)\underset{\alpha\geqslant 0}{\Rightarrow}\alpha\cdot x\leqslant\alpha\cdot\sup(A)$ giving

$$\sup(\{\alpha\cdot x|x\in A|\})\leqslant\alpha\cdot\sup(A) \tag{17.4}$$

Assume now that $\sup(\{\alpha\cdot x|x\in A|\})<\alpha\cdot\sup(A)\Rightarrow\frac{1}{\alpha}\cdot\sup(\{\alpha\cdot x|x\in A\})<\sup(A)$ hence by the definition of the supremum there exists a $z\in A$ such that $\frac{1}{\alpha}\cdot\sup(\{\alpha\cdot x|x\in A\})<z\leqslant\sup(A)\Rightarrow\sup(\{\alpha\cdot x|x\in A\})<\alpha\cdot z$ a contradiction as $\alpha\cdot z\in\{\alpha\cdot x|x\in A\}$. So we must have that $\alpha\cdot\sup(A)\leqslant\sup(\{\alpha\cdot x|x\in A|\})$ which together with 17.4 gives

$$\sup(\{\alpha\cdot x|x\in A|\})=\alpha\cdot\sup(A)$$

4. We must consider two cases for α

$\alpha=0$. In this cases $\{\alpha\cdot x|x\in A\}=\{0\}$ and thus $\inf(\{\alpha\cdot x|x\in A\})=\inf(\{0\})=0=0\cdot\inf A)=\alpha\cdot\inf(A)$

$0\neq\alpha$. $\forall x\in A$ we have have $\inf(A)\leqslant x\underset{\alpha\geqslant 0}{\Rightarrow}\alpha\cdot\inf(A)\leqslant\alpha\cdot x$ giving

$$\alpha\cdot\inf(A)\leqslant\inf(\{\alpha\cdot x|x\in A|\}) \tag{17.5}$$

Assume now that $\alpha \cdot \inf(A) < \inf(\{\alpha \cdot x | x \in A|\}) < \Rightarrow \inf(A) < \frac{1}{\alpha} \cdot \inf(\{\alpha \cdot x | x \in A\})$ hence by the definition of the infinum there exists a $z \in A$ such that $\inf(A) \leqslant z < \frac{1}{\alpha} \cdot \inf(\{\alpha \cdot x | x \in A\}) \Rightarrow \alpha \cdot z < \inf(\{\alpha \cdot x | x \in A\})$ a contradiction as $\alpha \cdot z \in \{\alpha \cdot x | x \in A\}$. So we must have that $\inf(\{\alpha \cdot x | x \in A|\}) \leqslant \inf(A)$ which together with 17.5gives

$$\inf(\{\alpha \cdot x | x \in A|\}) = \alpha \cdot \inf(A)$$

□

17.2 Topology on $\bar{\mathbb{R}}$

Definition 17.31. *Let $x \in \bar{\mathbb{R}}$ then $|x| = \left\{ \begin{smallmatrix} x \text{ if } 0 \leqslant x \\ -x \text{ if } x < 0 \end{smallmatrix} \right.$, $||$ is called the absolute value on the extended reals.*

Note 17.32. From the definition it is clear that $x \leqslant |x|$

The absolute value on the extended reals has similar properties as the absolute value on the real numbers. However we must take in account that the sum is not always defined.

Theorem 17.33. *The absolute value has the following properties*

1. *$\forall x \in \mathbb{R}$ we have that $|x| = |x|_r$ (where we use the special notation $||_r$ to indicate the absolute value on the real numbers (see 8.65)*
2. *$|x| = 0 \Leftrightarrow x = 0$*
3. *$\forall x, y \in \bar{\mathbb{R}}$ we have $|x \cdot y| = |x| \cdot |y|$*
4. *$\forall x, y \in \bar{\mathbb{R}}$ such that $x + y$ is well defined we have that $|x + y| \leqslant |x| + |y|$*

Proof.

1. If $x \in \mathbb{R}$ then $|x| = \left\{ \begin{smallmatrix} x \text{ if } x \geqslant 0 \\ -x \text{ if } x < 0 \end{smallmatrix} \right. \underset{8.65}{=} |x|_r$
2. First $|0| \underset{0 \leqslant 0}{=} 0$ by definition, second if $|x| = 0$ then if $x \neq 0$ we have either $x < 0 \Rightarrow 0 = |x| = -x \neq 0$ a contradiction or $0 < x \Rightarrow 0 = |x| = x \neq 0$ again a contradiction so we must have $x = 0$
3. We must consider the following cases for x, y

   $0 \leqslant x \wedge 0 \leqslant y$. then by 17.18 we have that $0 \leqslant x \cdot y$ so that $|x \cdot y| = x \cdot y = |x| \cdot |y|$

   $0 \leqslant x \wedge y \leqslant 0$. then by 17.18 we have that $x \cdot y \leqslant 0$ so that $|x \cdot y| = -(x \cdot y) = x \cdot (-y) = |x| \cdot |y|$

   $x \leqslant 0 \wedge 0 \leqslant y$. then by 17.18 we have that $x \cdot y \leqslant 0$ so that $|x \cdot y| = -(x \cdot y) = (-x) \cdot y = |x| \cdot |y|$

   $x \leqslant 0 \wedge y \leqslant 0$. then by 17.18 we have that $0 \leqslant x \cdot y$ so that $|x \cdot y| = x \cdot y = (-|x|) \cdot (-|y|) = (-1) \cdot (-1) \cdot |x| \cdot |y| = |x| \cdot |y|$

4. We have to consider the following (valid) cases for x, y

    $x \in \mathbb{R} \wedge y \in \mathbb{R}$. then $|x+y| = |x+y|_r \leqslant |x_r| + |y|_r = |x| + |y|$

    $x \in \mathbb{R} \wedge y = \infty$. then $|x+y| = |\infty| = \infty \leqslant \infty = |x| + \infty = |x| + |\infty|$

    $x \in \mathbb{R} \wedge y = -\infty$. then $|x+y| = |-\infty| = \infty \leqslant \infty = |x| + \infty = |x| + |y|$

    $x = \infty \wedge y \in \mathbb{R}$. then $|x+y| = |\infty| = \infty \leqslant \infty = \infty + |y| = |x| + |y|$

    $x = \infty \wedge y = \infty$. then $|x+y| = |\infty| = \infty \leqslant \infty = \infty + \infty = |x| + |y|$

    $x = -\infty \wedge y \in \mathbb{R}$. then $|x+y| = |\infty| = \infty \leqslant \infty = \infty + |y| = |x| + |y|$

    $x = -\infty \wedge y = -\infty$. then $|x+y| = |-\infty| = \infty \leqslant \infty = |x| + |y|$ □

To define the topology on $\bar{\mathbb{R}}$ we first define some sets of sets that we use to create a generating basis for the topology (see 12.28).

Notation 17.34. *In what follows $\varepsilon > 0$ means $\varepsilon \in \mathbb{R} \wedge \varepsilon > 0$*

Definition 17.35. *We define the following sets*

1. $\mathcal{B}_{\mathbb{R}} = \{]x-\varepsilon, x+\varepsilon[| x \in \mathbb{R} \wedge \varepsilon \in \mathbb{R} \vdash \varepsilon > 0\}$ *(the set of open balls in $\mathbb{R}$ using the metric defined by the norm $||$)*
2. $\mathcal{B}_{\infty} = \{]x, \infty] | x \in \mathbb{R}\}$
3. $\mathcal{B}_{-\infty} = \{[-\infty, x[| x \in \mathbb{R}\}$
4. $\mathcal{B}_{\bar{\mathbb{R}}} = \mathcal{B}_{\mathbb{R}} \bigcup \mathcal{B}_{\infty} \bigcup \mathcal{B}_{-\infty} \subseteq \mathcal{P}(\bar{\mathbb{R}})$

Theorem 17.36. *$\mathcal{B}_{\bar{\mathbb{R}}}$ forms a generating basis for a topology $\mathcal{T}_{\bar{\mathbb{R}}}$ on $\bar{\mathbb{R}}$. Further if $\mathcal{T}_{\mathbb{R}}$ is the canoncial topology on $\mathbb{R}$ (defined by the norm $||$) then $\mathcal{T}_{\mathbb{R}} \subseteq \mathcal{T}_{\bar{\mathbb{R}}}$ and $\mathcal{T}_{\mathbb{R}}$ is the subspace topology on $\mathbb{R}$ induced by $\mathcal{T}_{\bar{\mathbb{R}}}$, in other words $\mathcal{T}_{\mathbb{R}} = \{U \bigcap \mathbb{R} | U \in \mathcal{T}_{\bar{\mathbb{R}}}\}$*

Proof. We have

1. $\forall x \in \bar{\mathbb{R}}$ we have either

    $x \in \mathbb{R}$. then $x \in]x-1, x+1[\in \mathcal{B}_{\bar{\mathbb{R}}}$

    $x = \infty$. then $\infty \in]0, \infty] \in \mathcal{B}_{\bar{\mathbb{R}}}$

    $x = -\infty$. then $-\infty \in [-\infty, 0[\in \mathcal{B}_{\bar{\mathbb{R}}}$

    which proves that $\forall x \in \bar{\mathbb{R}}$ there exists a $B \in \mathcal{B}_{\bar{\mathbb{R}}}$ such that $x \in B$

2. If $A, B \in \mathcal{B}_{\bar{\mathbb{R}}}$ then we have either

    $A, B \in \mathcal{B}_{\mathbb{R}}$. then $A =]x-\varepsilon, x+\varepsilon[$, $B =]y-\delta, y+\delta[$, if $z \in A \bigcap B$ then we have $z \in \mathbb{R} \wedge x-\varepsilon < z < x+\varepsilon \wedge y-\delta < z < y+\delta$ so that $\rho = \min(z-x+\varepsilon, z-y+\delta, x+\varepsilon-z, y+\delta-z) > 0$, then we have $-z+x-\varepsilon < -\rho \Rightarrow x-\varepsilon < z-\rho$, $-z+y-\delta < -\rho \Rightarrow y-\delta < z-\rho$, $z+\rho < x+\varepsilon$ and $z+\rho < y+\delta$ proving that $]z-\rho, z+\rho[\subseteq A, B$. So there exists a $C \in \mathcal{B}_{\bar{\mathbb{R}}}$ with $z \in C \subseteq A \bigcap B$

    $A \in \mathcal{B}_{\infty} \wedge B \in \mathcal{B}_{\mathbb{R}}$. then $A =]x, \infty]$ and $B =]y-\delta, y+\delta[$. if $z \in A \bigcap B$ then we have $z \in \mathbb{R} \wedge x < z \wedge y-\delta < z < y+\delta$ so that $\rho = \min(z-x, z-y+\delta, y+\delta-z) > 0$, then we have $x-z < -\rho \Rightarrow x < z-\delta$, $-z+y-\delta < -\rho$ and $z+\rho < y+\delta$ proving that $]z-\rho, z+\rho[\subseteq A, B$. So there exists a $C \in \mathcal{B}_{\bar{\mathbb{R}}}$ with $z \in C \subseteq A \bigcap B$

$\boldsymbol{A \in \mathcal{B}_{-\infty} \wedge B \in \mathcal{B}_{\mathbb{R}}}$**.** then $A = [-\infty, x[$ and $B =]y - \delta, y + \delta[$, if $z \in A \bigcap B$ then $z \in \mathbb{R} \wedge z < x$, $y - \delta < z < y + \delta$ so that $\rho = \min(x - z, z - y + \delta, y + \delta - z) > 0$, then we have $-z + y - \delta < -\rho \Rightarrow y - \delta < z - \rho$, $z + \rho < x$ and $z + \rho < y + \delta$ proving that $]z - \rho, z + \rho[\subseteq A, B$. So there exists a $C \in \mathcal{B}_{\bar{\mathbb{R}}}$ with $z \in C \subseteq A \bigcap B$

$\boldsymbol{A \in \mathcal{B}_{\mathbb{R}} \wedge B \in \mathcal{B}_{\infty}}$**.** this is equal to the case $A \in \mathcal{B}_{\infty} \wedge B \in \mathcal{B}_{\mathbb{R}}$ by interchanging A and B

$\boldsymbol{A \in \mathcal{B}_{\mathbb{R}} \wedge B \in \mathcal{B}_{-\infty}}$**.** this is equal to the case $A \in \mathcal{B}_{-\infty} \wedge B \in \mathcal{B}_{\mathbb{R}}$ by interchanging A and B

$\boldsymbol{A \in \mathcal{B}_{\infty} \wedge B \in \mathcal{B}_{\infty}}$**.** then $A =]x, \infty]$ and $B =]y, \infty]$, if $z \in A \bigcap B$ then $x < z \wedge y < z \Rightarrow \max(x, y) < z \Rightarrow z \in]\max(x, y), \infty]$ and as $x, y \leqslant \max(x, y)$ we have $]\max(x, y), \infty] \subseteq A, B$. So there exists a $C \in \mathcal{B}_{\bar{\mathbb{R}}}$ with $z \in C \subseteq A \bigcap B$

$\boldsymbol{A \in \mathcal{B}_{\infty} \wedge B \in \mathcal{B}_{-\infty}}$**.** then $A =]x, \infty]$ and $B = [-\infty, y[$, if $z \in A \bigcap B$ then $x < z \wedge z < y \Rightarrow z \in \mathbb{R}$, take then $\rho = \min(z - x, y - z) > 0$, then we have $-z + x < -\rho \Rightarrow x < z + \rho$ and $z + \rho < y$ proving that $]z - \rho, z + \rho[\subseteq A, B$. So there exists a $C \in \mathcal{B}_{\bar{\mathbb{R}}}$ with $z \in C \subseteq A \bigcap B$

$\boldsymbol{A \in \mathcal{B}_{-\infty} \wedge B \in \mathcal{B}_{\infty}}$**.** this is equal to the case $A \in \mathcal{B}_{\infty} \wedge B \in \mathcal{B}_{-\infty}$ by interchanging A and B

$\boldsymbol{A \in \mathcal{B}_{-\infty} \wedge B \in \mathcal{B}_{-\infty}}$**.** then $A = [-\infty, x[$ and $B = [-\infty, y[$, if $z \in A \bigcap B$ then $z < y \wedge z < x \Rightarrow z < \min(x, y) \Rightarrow z \in [-\infty . \min(x, y)[$ and as $\min(x, y) \leqslant x, y$ we have $[-\infty, \min(x, y)[\subseteq A, B$. So there exists a $C \in \mathcal{B}_{\bar{\mathbb{R}}}$ with $z \in C \subseteq A \bigcap B$

This proves that $\forall A, B \in \mathcal{B}_{\bar{\mathbb{R}}}$ we have if $z \in A \bigcap B$ then there exists a $C \in \mathcal{B}_{\bar{\mathbb{R}}}$ such that $z \in C \subseteq A \bigcap B$.

Using the definition of a generating basis for a topology (see 12.28) we have by (1) and (2) that $\mathcal{B}_{\bar{\mathbb{R}}}$ is the generating basis for a topology on $\bar{\mathbb{R}}$, we call this topology $\mathcal{T}_{\bar{\mathbb{R}}}$, then we we have $\mathcal{T}_{\bar{\mathbb{R}}} \underset{12.28}{=} \{U \subseteq \bar{\mathbb{R}} | \forall x \in U \vDash \exists B \in \mathcal{B}_{\bar{\mathbb{R}}} \text{ such that } x \in B \subseteq U\}$. As $\mathcal{T}_{\mathbb{R}} = \{U \subseteq \mathbb{R} | \forall x \in U \vDash \exists B \in \mathcal{B}_{\mathbb{R}}\}$ we have as $\mathcal{B}_{\mathbb{R}} \subseteq \mathcal{B}_{\bar{\mathbb{R}}}$ and $\mathbb{R} \subseteq \bar{\mathbb{R}}$ proving that

$$\mathcal{T}_{\mathbb{R}} \subseteq \mathcal{T}_{\bar{\mathbb{R}}} \tag{17.6}$$

Finally if $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ then if $x \in U \bigcap \mathbb{R}$ there exists a $B' \in \mathcal{B}_{\bar{\mathbb{R}}}$ with $x \in B' \subseteq U$ we have either

$\boldsymbol{B' =]z - \varepsilon, z + \varepsilon[}$**.** then $B' \in \mathcal{B}_{\mathbb{R}} \underset{\text{take } B = B'}{\Rightarrow} \exists B \in \mathcal{B}_{\mathbb{R}} \vDash x \in B \subseteq U$

$\boldsymbol{B' =]z, \infty]}$**.** then as $x \in \mathbb{R}$ we have $x \neq \infty$ so that $x \in]z, \infty[\Rightarrow z < x$, then $\varepsilon = x - z > 0$ proving that $x \in]x - \varepsilon, x + \varepsilon[=]z, z + \varepsilon[\subseteq]z, \infty[\subseteq B' \subseteq U \underset{\text{take } B =]x - \varepsilon, x + \varepsilon[}{\Rightarrow} \exists B \in \mathcal{B}_{\mathbb{R}}$ with $x \in B \subseteq U$

$\boldsymbol{B' = [-\infty, z[}$**.** then as $x \in \mathbb{R}$ we have $x \neq -\infty$ so that $x \in]-\infty, z[\Rightarrow x < z$, then $\varepsilon = z - x > 0$ proving that $x \in]x - \varepsilon, x + \varepsilon[\subseteq]x - \varepsilon, z[\subseteq]-\infty, z[\subseteq B' \underset{\text{take } B =]x - \varepsilon, x + \varepsilon[}{\Rightarrow} \exists B \in \mathcal{B}_{\mathbb{R}}$ with $x \in B \subseteq U$

proving that $\forall x \in U \bigcap \mathbb{R}$ we have $\exists B \in \mathcal{B}_{\mathbb{R}}$ such that $x \in B \subseteq U \bigcap \mathbb{R}$ proving that $U \bigcap \mathbb{R} \in \mathcal{T}_{\mathbb{R}}$ and thus that

$$\left\{U \bigcap \mathbb{R} | U \in \mathcal{T}_{\bar{\mathbb{R}}}\right\} \subseteq \mathcal{T}_{\mathbb{R}} \tag{17.7}$$

If $U \in \mathcal{T}_{\mathbb{R}}$ then $U \subseteq \mathbb{R} \Rightarrow U = U \bigcap \mathbb{R}$ and as by 17.6 we have $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ we have that $U \in \{U \bigcap \mathbb{R} | U \in \mathcal{T}_{\bar{\mathbb{R}}}\}$ proving that $\mathcal{T}_{\mathbb{R}} \subseteq \{U \bigcap \mathbb{R} | U \in \mathcal{T}_{\bar{\mathbb{R}}}\}$ which together with 17.7 proves that $\mathcal{T}_{\mathbb{R}}$ is the subspace topology of $\mathcal{T}_{\bar{\mathbb{R}}}$ on $\mathbb{R}$. □

Theorem 17.37. *We have the following closed sets in $\mathcal{T}_{\bar{\mathbb{R}}}$ where $x \in \mathbb{R}$*

1. $[-\infty, x]$
2. $[x, \infty]$
3. $\{x\}$
4. $[-\infty, \infty]$
5. $\{-\infty, \infty\}$
6. $\{-\infty\}$
7. $\{\infty\}$

Proof.

1. As $y \in \bar{\mathbb{R}} \backslash [-\infty, x] \Leftrightarrow y \in \neg(-\infty \leqslant y \wedge y \leqslant x) \Leftrightarrow (y < -\infty \vee x < y) \Leftrightarrow x < y \Leftrightarrow y \in]x, \infty]$ we have that $\bar{\mathbb{R}} \backslash [-\infty, x] =]x, \infty]$ by definition a open set
2. As $y \in \bar{\mathbb{R}} \backslash [x, \infty] \Leftrightarrow y \in \neg(x \leqslant y \wedge y \leqslant \infty) \Leftrightarrow (y < x \vee \infty < y) \Leftrightarrow y < x \Leftrightarrow y \in [-\infty, x[$ we have that $\bar{\mathbb{R}} \backslash [x, \infty] = [-\infty, x[$ by definiton a open set.
3. $\{x\} = [-\infty, x] \bigcap [x, \infty]$ by (1) and (2) a intersection of closed sets and thus closed
4. This is trivial as $\bar{\mathbb{R}} = [-\infty, \infty]$
5. As $\bar{\mathbb{R}} \backslash \{-\infty, \infty\} = \mathbb{R} \in \mathcal{T}_{\mathbb{R}} \subseteq \mathcal{T}_{\bar{\mathbb{R}}}$ we have that $\bar{\mathbb{R}} \backslash \{-\infty, \infty\}$ is open so that $\{-\infty, \infty\}$ is closed.
6. This follows from $\{-\infty\} = \{-\infty, \infty\} \bigcap [-\infty, 0]$ a intersection of closed sets.
7. This follows from $\{\infty\} = \{-\infty, \infty\} \bigcap [0, \infty]$ a intersection of closed sets. □

We have the following characterization of open sets in $\bar{\mathbb{R}}$.

Theorem 17.38. *Let $x \in \bar{\mathbb{R}}$ and $x \in U$ a open set then we have*

1. *If $x \in \mathbb{R}$ then there exists a $\varepsilon > 0$ such that $x \in]x - \varepsilon, x + \varepsilon[\subseteq U$*
2. *If $x = \infty$ then there exists a $\varepsilon > 0$ such that $x \in]\varepsilon, \infty] \subseteq U$*
3. *If $x = -\infty$ then there exists a $\varepsilon > 0$ such that $x \in [-\infty, -\varepsilon[\subseteq U$*

Proof. Let $x \in \bar{\mathbb{R}}$ and $x \in U$ a open set then

1. If $x \in \mathbb{R}$ we have as $\mathcal{B}_{\bar{\mathbb{R}}}$ is a basis for $\mathcal{T}_{\bar{\mathbb{R}}}$ the following possible cases

   $\boldsymbol{\exists \delta > 0, y \in \mathbb{R} \vdash x \in]y - \delta, y + \delta[.}$ then $y - \delta < x < y + \delta$ take then $\varepsilon = \min(x - (y - \delta), (y + \delta) - x) > 0$. Then $x - \varepsilon > x - (x - (y - \delta)) = y - \delta$ and $x + \varepsilon < x + (y + \delta) - x = y + \delta$ so that $x \in]x - \varepsilon, x + \varepsilon[\subseteq]y - \delta, y + \delta[\subseteq U$ proving that $x \in]x - \varepsilon, x + \varepsilon[\subseteq U$.

$\exists\delta\in\mathbb{R}\vdash x\in]\delta,\infty]$. then $\delta<x<\infty$, take then $\varepsilon=x-\delta>0$ then we have $x-\varepsilon=x-x+\delta=\delta$ and $x+\varepsilon<\infty$ so that $x\in]x-\varepsilon,x+\varepsilon[\subseteq]\delta,\infty]\subseteq U$

$\exists\delta\in\mathbb{R}\vdash x\in[-\infty,\delta[$. then $-\infty<x<\delta$, take then $\varepsilon=\delta-x>0$ then we have $x+\varepsilon<\delta$ and $-\infty<x-\varepsilon$ so that $x\in]x-\varepsilon,x+\varepsilon[\subseteq[-\infty,\delta[\subseteq U$

2. If $x=\infty$ then we have as $\mathcal{B}_{\bar{\mathbb{R}}}$ is a basis for $\mathcal{T}_{\bar{\mathbb{R}}}$ the following possible cases

$\exists\delta>0,y\in\mathbb{R}\vdash x\in]y-\delta,y+\delta[$. then $\infty<y+\delta<\infty$ a contradiction so this will never happen.

$\exists\delta\in\mathbb{R}\vdash x\in]\delta,\infty]$. then as $1<\infty=x\leqslant\infty$ we have if we take $\varepsilon=\max(\delta,1)$ we have $\delta\leqslant\varepsilon$ so that $x\in]\varepsilon,\infty]\subseteq]\delta,\infty]\subseteq U$

$\exists\delta\in\mathbb{R}\vdash x\in[-\infty,\delta[$. then $\infty=x<\delta<\infty$ a contradiction so thiw will not apply

3. If $x=-\infty$ then we have as $\mathcal{B}_{\bar{\mathbb{R}}}$ is a basis for $\mathcal{T}_{\bar{\mathbb{R}}}$ the following possible cases

$\exists\delta>0,y\in\mathbb{R}\vdash x\in]y-\delta,y+\delta[$. then $-\infty<y-\delta<x=-\infty$ a contradiction.

$\exists\delta\in\mathbb{R}\vdash x\in]\delta,\infty]$. then $-\infty<\delta<x=-\infty$ a contradiction

$\exists\delta\in\mathbb{R}\vdash x\in[-\infty,\delta[$. then as $-\infty<-1$ we have if we take $\varepsilon=-\min(-1,\delta)>0$ that $-\varepsilon\leqslant\delta$ so that $x=-\infty\in[-\infty,-\varepsilon[\subseteq[-\infty,\delta[\subseteq U$ □

We show now that continuous functions to $\mathbb{R}$ are continuous in $\bar{\mathbb{R}}$

Theorem 17.39. *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $f:X\to\mathbb{R}$ a continuous function (using $\mathcal{T}_{\mathbb{R}}$) then $f:X\to\bar{\mathbb{R}}$ is continous in $\mathcal{T}_{\bar{\mathbb{R}}}$.*

Proof. Let $x\in X$ and $U\in f(x)$ then as $f(x)\in\mathbb{R}$ we have by the above theorem a $\varepsilon>0$ such that $]f(x)-\varepsilon,f(x+\varepsilon)[$. As $f:X\to\mathbb{R}$ is continuous there exists a $V\in\mathcal{T}$ such that $f(X)\subseteq]f(x)-\varepsilon,f(x0+\varepsilon)[\subseteq U$ proving that $f:X-\bar{\mathbb{R}}$ is continuous. □

Theorem 17.40. *If $U\in\mathcal{T}_{\bar{\mathbb{R}}}$ then if $\delta\in\mathbb{R}$ we have that $\delta+U=\{\delta+x|x\in U\}\in\mathcal{T}_{\bar{\mathbb{R}}}$ and if $\delta\neq 0$ we have $\delta\cdot U=\{\delta\cdot x|x\in U\}\in\mathcal{T}_{\bar{\mathbb{R}}}$*

Proof. If $x\in\delta+U$ then there exists a $u\in U$ such that $x=\delta+u$ we have then for u the following possibilities

$u\in\mathbb{R}$. then there exists a $\varepsilon>0$ such that $u\in]u-\varepsilon,u+\varepsilon[\subseteq U$ then $u-\varepsilon<u<u-\varepsilon\Rightarrow(u-\varepsilon)+\delta<u+\delta<(u+\varepsilon)+\delta$ proving that $x=u+\delta\in](u-\varepsilon)+\delta,(u+\varepsilon)+\delta[=](u+\delta)-\varepsilon,(u+\delta)+\varepsilon[$. Also if $y\in](u+\delta)-\varepsilon,(u+\delta)+\varepsilon[$ then $(u+\delta)-\varepsilon<y<(u+\delta)+\varepsilon\Rightarrow u-\varepsilon<y-\delta<u+\varepsilon\Rightarrow y-\delta\in]u-\varepsilon,u+\varepsilon[\subseteq U$ so that $y=\delta+(y-\delta)\in\delta+U$ proving that $x\in](u+\delta)-\varepsilon,(u+\delta)+\varepsilon[\subseteq\delta+U$. This proves that $\exists B\in\mathcal{B}_{\bar{\mathbb{R}}}$ such that $x\in B\subseteq\delta+U$

$u=\infty$. then there exists a $\varepsilon>0$ such that $u=\infty\in]\varepsilon,\infty]\subseteq U$, take now $]\varepsilon+\delta,\infty]$ then $x=\delta+\infty=\infty\in]\varepsilon+\delta,\infty]$. If now $y\in]\varepsilon+\delta,\infty]$ then we have either $y=\infty=\delta+\infty=\delta+u\in\delta+U$ or $y\in\mathbb{R}$ so that $\varepsilon+\delta<y<\infty\Rightarrow\varepsilon<y-\delta<\infty\Rightarrow y-\delta\subset]\varepsilon,\infty]\subseteq U\Rightarrow y=\delta+y-\delta\in\delta+U$ proving that $x\in]\varepsilon+\delta,\infty]\subseteq\delta+U$. This proves that $\exists B\in\mathcal{B}_{\bar{\mathbb{R}}}$ such that $x\in B\subseteq\delta+U$.

$u = -\infty$. then there exists a $\varepsilon > 0$ such that $u = -\infty \in [-\infty, -\varepsilon[\subseteq U$, take now $[-\infty, -\varepsilon + \delta[$ then $x = \delta + (-\infty) = -\infty \in [-\infty, -\varepsilon + \delta[$. If now $y \in [-\infty, -\varepsilon + \delta[$ the we have either $y = -\infty = \delta + (-\infty) = \delta + u \in \delta + U$ or $y \in \mathbb{R}$ so that $-\infty < y < -\varepsilon + \delta \Rightarrow -\infty < y - \delta < -\varepsilon \Rightarrow y - \delta \in [-\infty, -\varepsilon[\subseteq U \Rightarrow y = \delta + (y - \delta) \in \delta + U$ proving that $x \in [-\infty, -\varepsilon + \delta[\subseteq U$. This proves that $\exists B \in \mathcal{B}_{\bar{\mathbb{R}}}$ such that $x \in B \subseteq U$

As in all possible cases we have $\forall x \in \delta + U$ that $\exists B \in \mathcal{B}_{\bar{\mathbb{R}}}$ such that $x \in B \subseteq \delta + U$ proving that $\delta + U$ is a open set.

If $x \in \delta \cdot U$ then there exists a $u \in U$ such that $x = \delta \cdot u$ we have then for u the following possibilities

$u \in \mathbb{R}$. then there exists a $\varepsilon > 0$ such that $u \in]u - \varepsilon, u + \varepsilon[\subseteq U$ so that $u - \varepsilon < u < u + \varepsilon$ consider then the following possibilities for $\delta \neq 0$:

$0 < \delta$. then $\delta \cdot u - \delta \cdot \varepsilon < \delta \cdot u < \delta \cdot u + \delta \cdot \varepsilon$ so that $x = \delta \cdot u \in]\delta \cdot u - \delta \cdot \varepsilon, \delta \cdot u + \delta \cdot \varepsilon[=]u \cdot \delta - \rho, u \cdot \delta + \rho[$ where $\rho = \delta \cdot \varepsilon > 0$. If now $y \in]u \cdot \delta - \rho, u \cdot \delta + \rho[\Rightarrow u \cdot \delta - \rho < y < u \cdot \delta + \rho \Rightarrow \frac{u \cdot \delta}{\delta} - \frac{\rho}{\delta} < \frac{y}{\delta} < \frac{u \cdot \delta}{\delta} + \frac{\rho}{\delta} \Rightarrow u - \varepsilon < \frac{y}{\delta} < u + \varepsilon \Rightarrow \frac{y}{\delta} \in]u - \varepsilon, u + \varepsilon[\subseteq U \Rightarrow y = \delta \cdot \frac{y}{\delta} \in \delta \cdot U$ proving that $x \in]\delta \cdot u - \rho, \delta \cdot u + \rho[\subseteq \delta \cdot U$. This proves that $\exists B \in \mathcal{B}_{\bar{\mathbb{R}}}$ such that $x \in B \subseteq \delta \cdot U$

$\delta < 0$. then $\delta \cdot u + \delta \cdot \varepsilon < \delta \cdot u < \delta \cdot u - \delta \cdot \varepsilon \Rightarrow x = \delta \cdot u \in]\delta \cdot u + \delta \cdot \varepsilon, \delta \cdot u - \delta \cdot \varepsilon[=]\delta \cdot u - \rho, \delta \cdot u + \rho[$ where $\rho = -\delta \cdot \varepsilon > 0$. If now $y \in]u \cdot \delta - \rho, u \cdot \delta + \rho[\Rightarrow u \cdot \delta - \rho < y < u \cdot \delta + \rho \Rightarrow \frac{u \cdot \delta}{\delta} + \frac{\rho}{\delta} < \frac{y}{\delta} < \frac{u \cdot \delta}{y} - \frac{\rho}{\delta} \Rightarrow u - \varepsilon < \frac{y}{\delta} < u + \varepsilon \Rightarrow \frac{y}{\rho} \in]u - \varepsilon, u + \varepsilon[\subseteq U$ so that $y = \delta \cdot \frac{y}{\delta} \in \delta \cdot U$ proving that $x \in]u - \varepsilon, u + \varepsilon[\subseteq \delta \cdot U$. This proves that $\exists B \in \mathcal{B}_{\bar{\mathbb{R}}}$ such that $x \in B \subseteq \delta \cdot U$

$u = \infty$. then there exists a $\varepsilon > 0$ such that $u \in]\varepsilon, \infty] \subseteq U$, we have to consider the following cases for δ

$0 < \delta$. then $x = \delta \cdot u = \infty \in]\varepsilon \cdot \delta, \infty]$. If $y \in]\varepsilon \cdot \delta, \infty]$ then if $y = \infty = \delta \cdot u$ we have $y = \delta \cdot u \in \delta \cdot U$ and if $y \in \mathbb{R}$ we have that $\varepsilon \cdot \delta < y < \infty$ so that $\varepsilon < \frac{y}{\delta} < \infty \Rightarrow \frac{y}{\delta} \in]\varepsilon, \infty] \subseteq U \Rightarrow y = \delta \cdot \frac{y}{\delta} \in \delta \cdot U$. So we have proved that $x \in]\varepsilon \cdot \delta, \infty] \subseteq U$ and thus that $\exists B \in \mathcal{B}_{\bar{\mathbb{R}}}$ such that $x \in B \subseteq \delta \cdot U$

$\delta < 0$. then $x = \delta \cdot u = -\infty \in [-\infty, \varepsilon \cdot \delta[$. If $y \in [-\infty, \varepsilon \cdot \delta[$ then if $y = -\infty = \delta \cdot u$ we have $y = \delta \cdot \infty = \delta \cdot u \in \delta \cdot U$ and if $y \in \mathbb{R}$ we have that $-\infty < y < \varepsilon \cdot \delta$ so that $\varepsilon < \frac{y}{\delta} < \infty \Rightarrow \frac{y}{\delta} \in]\varepsilon, \infty] \subseteq U \Rightarrow y = \delta \cdot \frac{y}{\delta} \in \delta \cdot U$ So we have proved that $x \in]\varepsilon, \infty] \subseteq \delta \cdot U$ and thus that $\exists B \in \mathcal{B}_{\bar{\mathbb{R}}}$ such that $x \in B \subseteq \delta \cdot U$

$u = -\infty$. then there exists a $\varepsilon > 0$ such that $u \in [-\infty, -\varepsilon[\subseteq U$, we have to consider the following cases for δ

$0 < \delta$. then $x = \delta \cdot u = -\infty \in [-\infty, -\varepsilon \cdot \delta[$. If $y \in [-\infty, -\varepsilon \cdot \delta[$ then if $y = -\infty$ we have $y = \delta \cdot (-\infty) = \delta \cdot u \in \delta \cdot U$ and if $y \in \mathbb{R}$ we have that $-\infty < y < -\varepsilon \cdot \delta$ so that $-\infty < \frac{y}{\delta} < -\varepsilon \Rightarrow \frac{y}{\delta} \in [-\infty, -\varepsilon[\subseteq U \Rightarrow y = \delta \cdot \frac{y}{\delta} \in \delta \cdot U$. So we have proved that $x \in [-\infty, -\varepsilon \cdot \delta[\subseteq U$ and thus that $\exists B \in \mathcal{B}_{\bar{\mathbb{R}}}$ such that $x \in B \subseteq \delta \cdot U$

$\delta < 0$. then $x = \delta \cdot u = \infty \in]\varepsilon \cdot \delta, \infty]$. If $y \in]\varepsilon \cdot \delta, \infty]$ then if $y = \infty$ we have $y = \delta \cdot -\infty = \delta \cdot u \in \delta \cdot U$ and if $y \in \mathbb{R}$ we have that $\varepsilon \cdot \delta < y < \infty$ so that $-\infty < \frac{y}{\delta} < \varepsilon \Rightarrow \frac{y}{\delta} \in [-\infty, \varepsilon[\subseteq U \Rightarrow y = \delta \cdot \frac{y}{\delta} \in \delta \cdot U$ So we have proved that $x \in]\varepsilon \cdot \delta, \infty] \subseteq \delta \cdot U$ and thus that $\exists B \in \mathcal{B}_{\bar{\mathbb{R}}}$ such that $x \in B \subseteq \delta \cdot U$

As in all possible cases we have $\forall x \in \delta \cdot U$ that $\exists B \in \mathcal{B}_{\bar{\mathbb{R}}}$ such that $x \in B \subseteq \delta \cdot U$ which proves that $\delta \cdot U$ is open. □

Note 17.41. The restrictions in the above theorem are need for example is $U =]-1, 1[$ a open set then $0 \cdot U = 0 \cdot]-1, 1[= \{0\}$ a closed set. Also $\infty + U = \infty +]-1, 1[= \{\infty\}$ a closed set.

Theorem 17.42. *$\mathcal{T}_{\bar{\mathbb{R}}}$ is Hausdorff (see 12.218)*

Proof. Let $x_1, x_2 \in \bar{\mathbb{R}}$ with $x_1 \neq x_2$ then we have to check the following cases

$x_1 \in \mathbb{R} \wedge x_2 \in \mathbb{R}$. then we have either

$x_1 < x_2$. Take then $\delta = \frac{x_2 - x_1}{2} > 0$ then $x_1 \in]x_1 - \delta, x_2 + \delta[\in \mathcal{T}_{\bar{\mathbb{R}}}$, $x_2 \in]x_2 - \delta, x_2 + \delta[\in \mathcal{T}_{\bar{\mathbb{R}}}$, as $]x_1 - \delta, x_1 + \delta[=]x_1 - \delta, \frac{x_1 + x_2}{2}[$, $]x_2 - \delta, x_2 + \delta[=]\frac{x_1 + x_2}{2}, x_2 + \delta[$ we have $\emptyset =]x_1 - \delta, x_1 + \delta[\bigcap]x_2 - \delta, x_2 + \delta[$.

$x_2 < x_1$. Take then $\delta = \frac{x_1 - x_2}{2} > 0$ then $x_1 \in]x_1 - \delta, x_2 + \delta[\in \mathcal{T}_{\bar{\mathbb{R}}}$, $x_2 \in]x_2 - \delta, x_2 + \delta[\in \mathcal{T}_{\bar{\mathbb{R}}}$, as $]x_2 - \delta, x_2 + \delta[=]x_2 - \delta, \frac{x_1 + x_2}{2}[$, $]x_1 - \delta, x_1 + \delta[=]\frac{x_1 + x_2}{2}, x_1 + \delta[$ we have $\emptyset =]x_1 - \delta, x_1 + \delta[\bigcap]x_2 - \delta, x_2 + \delta[$.

$x_1 \in \mathbb{R} \wedge x_2 = \infty$. then $x_1 \in]x_1 - 1, x_1 + 1[\in \mathcal{T}_{\bar{\mathbb{R}}}$ and $\infty \in]x_1 + 1, \infty] \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $]x_1 - 1, x_1 + 1[\bigcap]x_1 + 1, \infty] = \emptyset$

$x_1 \in \mathbb{R} \wedge x_2 = -\infty$. then $x_1 \in]x_1 - 1, x_1 + 1[\in \mathcal{T}_{\bar{\mathbb{R}}}$ and $-\infty \in [-\infty, x_1 - 1[\in \mathcal{T}_{\bar{\mathbb{R}}}$ with $]x_1 - 1, x_1 + 1[\bigcap [-\infty, x_1 - 1[= \emptyset$

$x_1 = \infty \wedge x_2 \in \mathbb{R}$. then $\infty \in]x_2 + 1, \infty] \in \mathcal{T}_{\bar{\mathbb{R}}}$ and $x_2 \in]x_2 - 1, x_2 + 1[\in \mathcal{T}_{\bar{\mathbb{R}}}$ with $]x_2 + 1, \infty] \bigcap]x_2 - 1, x_2 + 1[= \emptyset$

$x_1 = \infty \wedge x_2 = -\infty$. then $\infty \in]0, \infty] \in \mathcal{T}_{\bar{\mathbb{R}}}$ and $-\infty \in [-\infty, 0[$ with $]0, \infty] \bigcap [-\infty, 0[= \emptyset$

$x_1 = -\infty \wedge x_2 \in \mathbb{R}$. then $-\infty \in [-\infty, x_2 - 1[\in \mathcal{T}_{\bar{\mathbb{R}}}$ and $x_2 \in]x_2 - 1, x_2 + 1[\in \mathcal{T}_{\bar{\mathbb{R}}}$ with $[-\infty, x_2 - 1[\bigcap]x_2 - 1, x_2 + 1[= \emptyset$

$x_1 = -\infty \wedge x_2 = \infty$. then $-\infty \in [-\infty, 0[\in \mathcal{T}_{\bar{\mathbb{R}}}$ and $\infty \in]0, \infty] \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $[-\infty, 0[\bigcap]0, \infty] = \emptyset$

So in all cases there exists a $V_1, V_2 \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $x \in V_1 \wedge x \in V_2 \wedge V_1 \bigcap V_2 = \emptyset$ proving that $\mathcal{T}_{\bar{\mathbb{R}}}$ is Hausdorff. □

17.2.1 Limit of functions in $\bar{\mathbb{R}}$

Definition 17.43. *Let $A \subseteq \bar{\mathbb{R}}$, $x \in \bar{\mathbb{R}}$ then we have*

1. *x is a **left limit point** of A iff $\forall V \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $x \in V$ there exists a $y \in A \bigcap V$ with $y < x$*

2. *y is a **right limit point** of A iff $\forall V \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $x \in V$ there exists a $y \in A \bigcap V$ with $x < y$*

3. *y is a **limit point** of A iff $\forall V \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $x \in V$ there exists a $y \in A \bigcap V$ with $x \neq y$*

Lemma 17.44. *Let $A \subseteq \bar{\mathbb{R}}$ then we have*

1. *If x is a left limit point of A then $-\infty < x$*

2. *If x is a right limit point of A then $x < \infty$*

Proof.

1. Let x be a left limit point of A and assume that $x = -\infty$ then for the open set $[-\infty, 1[$ we have that $x \in [-\infty, 1[$ and thus by definiton there exists $y \in A \bigcap [-\infty, 1[$ with $y < x = -\infty$ a ontradiction.

2. Let x be a right limit point of A and assume that $x = \infty$ then for the open set $]1, \infty]$ we have that $x \in]1, \infty]$ and thus by definition there exists $y \in A \bigcap]1, \infty]$ with $\infty = x < y$ a contradiction. □

Remark 17.45. If $A \subseteq \bar{\mathbb{R}}$ then ∞ can not be a right limit point of A and $-\infty$ can not be a left limit point of A

Proof. This follows as there does not exists a $x \in \bar{\mathbb{R}}$ with $x < -\infty$ or $\infty < x$ □

Definition 17.46. *Let $A, B \subseteq \bar{\mathbb{R}}$, $f: A \rightarrow B$ a function $x_0, L \in \bar{\mathbb{R}}$ then we say that $\lim_{x \uparrow x_0} f(x) = L$ if and only if x_0 is a left limit point of A and $\forall U \in \mathcal{T}_{\bar{\mathbb{R}}}$ (see 17.36) with $L \in U$ there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $x_0 \in V$ so that $\forall x \in V \bigcap A$ with $x < x_0$ we have $f(x) \in U$. L is called the **left limit of f at x_0**.*

Definition 17.47. *Let $A, B \subseteq \bar{\mathbb{R}}$, $f: A \rightarrow B$ a function $x_0, L \in \bar{\mathbb{R}}$ then we say that $\lim_{x \downarrow x_0} f(x) = L$ if and only if x_0 is a right limit point of A and $\forall U \in \mathcal{T}_{\bar{\mathbb{R}}}$ (see 17.36) with $L \in U$ there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $x_0 \in V$ so that $\forall x \in V \bigcap A$ with $x_0 < x$ we have $f(x) \in U$. L is called the **right limit of f at x_0**.*

Definition 17.48. *Let $A, B \subseteq \bar{\mathbb{R}}$, $f: A \rightarrow B$ a function $x_0, L \in \bar{\mathbb{R}}$ then we say that $\lim_{x \rightarrow x_0} f(x) = L$ if and only if x_0 is a limit point of A and $\forall U \in \mathcal{T}_{\bar{\mathbb{R}}}$ (see 17.36) with $L \in U$ there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $x_0 \in V$ so that $\forall x \in V \bigcap A \vdash x \neq x_0$ we have $f(x) \in U$. L is called the **limit of f at x_0**.*

Note that the limitation of limits of functions to left limit points, right limit points or limit points is needed to have a unique limit (as is suggested in the notation for a limit). For example if $f:]0, 1[\rightarrow \bar{\mathbb{R}}$ is defined by $x \rightarrow 2 \cdot x$ so that 0 is not a left limit point of $]0, 1[$, then if we take $1, 2 \in \bar{\mathbb{R}}$ and $U_1, U_2 \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $1 \in U_1$, $2 \in U_2$ we have for $0 \in]-1, 1[\in \mathcal{T}_{\bar{\mathbb{R}}}$ that $\forall x \in V \bigcap A \vdash x < 0$ we have $f(x) \in U_1, U_2$ is satisfied vacuously. The following theorem shows that the limits as defined above have indeed unique limits and that the notation make sense.

Example 17.49. Let $A, B \subseteq \bar{\mathbb{R}}$, $f: A \rightarrow B$ a function, $x_0, L \in \bar{\mathbb{R}}$ then we have

1. If x_0 is a left limit point of A, $x_0 \in W$ a open set and $\forall x \in A \bigcap W \vdash x < x_0$ we have $f(x) = L$ then $\lim_{x \uparrow x_0} f(x) = L$
2. If x_0 is a a right limit point of A, $x_0 \in W$ a open set and $\forall x \in A \bigcap W \vdash x_0 < x$ we have $f(x) = L$ then $\lim_{x \downarrow x_0} f(x) = L$
3. If x_0 is a limit point of A, $x_0 \in W$ a open set and $\forall x \in A \bigcap W \vdash x \neq x_0$ we have $f(x) = L$ then $\lim_{x \rightarrow x_0} f(x) = L$

Proof.

1. Let U be a open set in $\bar{\mathbb{R}}$ such that $L \in U$ then if $x \in W \bigcap A$ with $x < x_0$ we have $f(x) = L \in U$ proving that $\lim_{x \uparrow x_0} f(x) = L$
2. Let U be a open set in $\bar{\mathbb{R}}$ such that $L \in U$ then if $x \in W \bigcap A$ with $x_0 < x$ we have $f(x) = L \in U$ proving that $\lim_{x \downarrow x_0} f(x) = L$
3. Let U be a open set in $\bar{\mathbb{R}}$ such that $L \in U$ then if $x \in W \bigcap A = A$ with $x \neq x_0$ we have $f(x) = L \in U$ proving that $\lim_{x \rightarrow x_0} f(x) = L$ □

Theorem 17.50. *Let $A, B \subseteq \bar{\mathbb{R}}$, $f: A \rightarrow B$ and $L_1, L_2, x_0 \in \bar{\mathbb{R}}$ then we have*

1. *If $\lim_{x \uparrow x_0} f(x) = L_1$ and $\lim_{x \uparrow x_0} f(x) = L_2$ then $L_1 = L_2$*
2. *If $\lim_{x \downarrow x_0} f(x) = L_1$ and $\lim_{x \downarrow x_0} f(x) = L_2$ then $L_1 = L_2$*
3. *If $\lim_{x \rightarrow x_0} f(x) = L_1$ and $\lim_{x \rightarrow x_0} f(x) = L_2$ then $L_1 = L_2$*

Proof. We prove this by contradiction

1. Assume that $L_1 \neq L_2$ then as $\mathcal{T}_{\bar{\mathbb{R}}}$ is Hausdorff (see 17.42) there exists U_1, $U_2 \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L_1 \in U_1 \wedge L_2 \in U_2 \wedge U_1 \bigcap U_2 = \emptyset$. Then $\exists V_1, V_2 \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $x_0 \in V_1 \wedge x_0 \in V_2 \wedge \forall x \in V_1 \bigcap A \vdash x < x_0$ we have $f(x) \in U_1 \wedge \forall x \in V_2 \bigcap A \vdash x < x_0$ we have $f(x) \in U_2$. As x_0 is a left limit point and $x_0 \in V_1 \bigcap V_2 \in \mathcal{T}_{\bar{\mathbb{R}}}$ there exists a $y \in (V_1 \bigcap V_2) \bigcap A$ with $y < x_0$ so that from the above we have $f(y) \in U_1 \bigcap U_2 = \emptyset$ a contradiction. So we must have $L_1 = L_2$.
2. Assume that $L_1 \neq L_2$ then as $\mathcal{T}_{\bar{\mathbb{R}}}$ is Hausdorff (see 17.42) there exists U_1, $U_2 \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L_1 \in U_1 \wedge L_2 \in U_2 \wedge U_1 \bigcap U_2 = \emptyset$. Then $\exists V_1, V_2 \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $x_0 \in V_1 \wedge x_0 \in V_2 \wedge \forall x \in V_1 \bigcap A \vdash x_0 < x$ we have $f(x) \in U_1 \wedge \forall x \in V_2 \bigcap A \vdash x_0 < x$ we have $f(x) \in U_2$. As x_0 is a right limit point and $x_0 \in V_1 \bigcap V_2 \in \mathcal{T}_{\bar{\mathbb{R}}}$ there exists a $y \in (V_1 \bigcap V_2) \bigcap A$ with $x_0 < y$ so that from the above we have $f(y) \in U_1 \bigcap U_2 = \emptyset$ a contradiction. So we must have $L_1 = L_2$.
3. Assume that $L_1 \neq L_2$ then as $\mathcal{T}_{\bar{\mathbb{R}}}$ is Hausdorff (see 17.42) there exists U_1, $U_2 \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L_1 \in U_1 \wedge L_2 \in U_2 \wedge U_1 \bigcap U_2 = \emptyset$. Then $\exists V_1, V_2 \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $x_0 \in V_1 \wedge x_0 \in V_2 \wedge \forall x \in V_1 \bigcap A \vdash x_0 \neq x$ we have $f(x) \in U_1 \wedge \forall x \in V_2 \bigcap A \vdash x_0 \neq x$ we have $f(x) \in U_2$. As x_0 is a limit point and $x_0 \in V_1 \bigcap V_2 \in \mathcal{T}_{\bar{\mathbb{R}}}$ there exists a $y \in (V_1 \bigcap V_2) \bigcap A$ with $x_0 \neq y$ so that from the above we have $f(y) \in U_1 \bigcap U_2 = \emptyset$ a contradiction. So we must have $L_1 = L_2$. □

Theorem 17.51. *Let $A, B \subseteq \bar{\mathbb{R}}$, $x_0 \in \bar{\mathbb{R}}$ then we have*

1. *If $\alpha \in \mathbb{R}$, $f: A \rightarrow B$ a function with a left limit $\lim_{x \uparrow x_0} f(x)$ then for $\alpha \cdot f: A \rightarrow C$ defined by $(\alpha \cdot f)(x) = \alpha \cdot f(x)$ where $(\alpha \cdot f)(A) \subseteq C$ we have $\lim_{x \uparrow x_0} (\alpha \cdot f)(x) = \alpha \cdot \lim_{x \uparrow x_0} f(x)$*
2. *If $\alpha \in \mathbb{R}$, $f: A \rightarrow B$ a function with a right limit $\lim_{x \downarrow x_0} f(x)$ then for $\alpha \cdot f: A \rightarrow C$ defined by $(\alpha \cdot f)(x) = \alpha \cdot f(x)$ where $(\alpha \cdot f)(A) \subseteq C$ we have $\lim_{x \downarrow x_0} (\alpha \cdot f)(x) = \alpha \cdot \lim_{x \downarrow x_0} f(x)$*
3. *If $\alpha \in \mathbb{R}$, $f: A \rightarrow B$ a function with a limit $\lim_{x \rightarrow x_0} f(x) = L$ then for $\alpha \cdot f: A \rightarrow C$ defined by $(\alpha \cdot f)(x) = \alpha \cdot f(x)$ where $(\alpha \cdot f)(A) \subseteq C$ we have $\lim_{x \rightarrow x_0} (\alpha \cdot f)(x) = \alpha \cdot \lim_{x \rightarrow x_0} f(x)$*
4. *If $f: A \rightarrow B$, $g: A \rightarrow B$ are functions with left limits $\lim_{x \uparrow x_0} f(x)$, $\lim_{x \uparrow x_0} g(x)$ such that $\forall x \in A$ we have that $f(x) + g(x)$ is defined and $\lim_{x \downarrow x_0} f(x) + \lim_{x \downarrow x_0} f(x)$ is defined (so not $\infty + (-\infty)$) then for $f + g: A \rightarrow C$ defined by $(f + g)(x) = f(x) + g(x)$ where $(f + g)(A) \subseteq C$ we have $\lim_{x \uparrow x_0} (f + g)(x) = \lim_{x \uparrow x_0} f(x) + \lim_{x \uparrow x_0} g(x)$*
5. *If $f: A \rightarrow B$, $g: A \rightarrow B$ are functions with left limits $\lim_{x \downarrow x_0} f(x)$, $\lim_{x \downarrow x_0} g(x)$ such that $\forall x \in A$ we have that $f(x) + g(x)$ is defined and $\lim_{x \uparrow x_0} f(x) + \lim_{x \uparrow x_0} f(x)$ is defined (so not $\infty + (-\infty)$) then for $f + g: A \rightarrow C$ defined by $(f + g)(x) = f(x) + g(x)$ where $(f + g)(A) \subseteq C$ we have $\lim_{x \downarrow x_0} (f + g)(x) = \lim_{x \downarrow x_0} f(x) + \lim_{x \downarrow x_0} g(x)$*
6. *If $f: A \rightarrow B$, $g: A \rightarrow B$ are functions with left limits $\lim_{x \uparrow x_0} f(x)$, $\lim_{x \uparrow x_0} g(x)$ such that $\forall x \in A$ we have that $f(x) + g(x)$ is defined and $\lim_{x \rightarrow x_0} f(x) + \lim_{x \rightarrow x_0} f(x)$ is defined (so not $\infty + (-\infty)$) then for $f + g: A \rightarrow C$ defined by $(f + g)(x) = f(x) + g(x)$ where $(f + g)(A) \subseteq C$ we have $\lim_{x \uparrow x_0} (f + g)(x) = \lim_{x \uparrow x_0} f(x) + \lim_{x \uparrow x_0} g(x)$*

Proof.

1. If $\alpha = 0$ then $\alpha \cdot f$ is the constant function 0 so that $\lim_{x \downarrow x_0} (\alpha \cdot f)(x) = 0 = 0 \cdot \lim_{x \downarrow x_0} f(x) = \alpha \cdot \lim_{x \downarrow x_0} f(x)$, proving the case of $a = 0$. So we are left with the case $\alpha \neq 0$. Let now U be a open set with $\alpha \cdot L \in U$ then $L \in \frac{1}{\alpha} \cdot U$ which is a open set (see 17.40). Using the definition of the right limit we have that x_0 is a right limit point of A and there exists a open set V containing x_0 such that $\forall x \in V \bigcap A$ with $x_0 < x$ we have $f(x) \in \frac{1}{\alpha} \cdot U \Rightarrow \alpha \cdot f(x) \in U$ proving that $\lim_{x \downarrow x_0} (\alpha \cdot f)(x) = \alpha \cdot L$.

2. If $\alpha = 0$ then $\alpha \cdot f$ is the constant function 0 so that $\lim_{x \uparrow x_0} (\alpha \cdot f)(x) = 0 = 0 \cdot \lim_{x \uparrow x_0} f(x) = \alpha \cdot \lim_{x \uparrow x_0} f(x)$, proving the case of $a = 0$. So we are left with the case $\alpha \neq 0$. Let now U be a open set with $\alpha \cdot L \in U$ then $L \in \frac{1}{\alpha} \cdot U$ which is a open set (see 17.40). Using the definition of the left limit we have that x_0 is a left limit point of A and there exists a open set V containing x_0 such that $\forall x \in V \bigcap A$ with $x_0 < x$ we have $f(x) \in \frac{1}{\alpha} \cdot U \Rightarrow \alpha \cdot f(x) \in U$ proving that $\lim_{x \uparrow x_0} (\alpha \cdot f)(x) = \alpha \cdot L$.

3. If $\alpha = 0$ then $\alpha \cdot f$ is the constant function 0 so that $\lim_{x \rightarrow x_0} (\alpha \cdot f)(x) = 0 = 0 \cdot \lim_{x \rightarrow x_0} f(x) = \alpha \cdot \lim_{x \rightarrow x_0} f(x)$, proving the case of $a = 0$. So we are left with the case $\alpha \neq 0$. Let now U be a open set with $\alpha \cdot L \in U$ then $L \in \frac{1}{\alpha} \cdot U$ which is a open set (see 17.40). Using the definition of the limit we have that x_0 is a limit point of A and there exists a open set V containing x_0 such that $\forall x \in V \bigcap A$ with $x_0 < x$ we have $f(x) \in \frac{1}{\alpha} \cdot U \Rightarrow \alpha \cdot f(x) \in U$ proving that $\lim_{x \rightarrow x_0} (\alpha \cdot f)(x) = \alpha \cdot L$.

4. Let U be a open set containing $L_1 + L_2$ then we have the following cases to consider for $L_1 + L_2$

   $L_1 + L_2 \in \mathbb{R}$. then we must have $L_1, L_2 \in \mathbb{R}$, If now $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in]L_1 + L_2 - \varepsilon, L_1 + L_2 - \varepsilon[\subseteq U$. From $L_1 \in \left] L_1 - \frac{\varepsilon}{2}, L_1 + \frac{\varepsilon}{2} \right[$, $L_2 \in \left] L_2 - \frac{\varepsilon}{2}, L_2 + \frac{\varepsilon}{2} \right[$ and the definition of right limits we have that x_0 is a right limit point of A and there exists open sets V_f, V_g with $x_0 \in V_f, V_g$ such that $\forall x \in V_f \bigcap A$ with $x_0 < x$ we have $f(x) \in \left] L_1 - \frac{\varepsilon}{2}, L_2 - \frac{\varepsilon}{2} \right[$ and $\forall x \in V_g \bigcap A$ with $x_0 < x$ we have $g(x) \in \left] L_2 - \frac{\varepsilon}{2}, L_2 + \frac{\varepsilon}{2} \right[$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x_0 < x$ then $L_1 - \frac{\varepsilon}{2} < f(x) < L_1 + \frac{\varepsilon}{2} \wedge L_2 - \frac{\varepsilon}{2} < g(x) < L_2 + \frac{\varepsilon}{2} \Rightarrow L_1 + L_2 - \varepsilon < f(x) + g(x) < L_1 + L_2 + \varepsilon \Rightarrow (f + g)(x) = f(x) + g(x) \in]L_1 + L_2 + \varepsilon, L_1 + L_2 - \varepsilon[\subseteq U$ which proves that $\lim_{x \downarrow x_0} (f + g)(x) = L_1 + L_2 = \lim_{x \downarrow x_0} f(x) + \lim_{x \downarrow x_0} g(x)$.

   $L_1 + L_2 = \infty$. then we have either

   $L_1, L_2 = \infty$. Let U be a open set with $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in]\delta, \infty] \subseteq U$ then as $L_1, L_2 = \infty$ we have $L_1, L_2 \in \left] \frac{\delta}{2}, \infty \right]$ and by the definition of right limits we have that x_0 is a right limit point of A and there exists open sets V_f,V_g containing x_0 such that $\forall x \in V_f \bigcap A$ with $x_0 < x$ we have $f(x) \in \left] \frac{\delta}{2}, \infty \right]$ and $\forall x \in V_g \bigcap A$ with $x_0 < x$ we have $g(x) \in \left] \frac{\delta}{2}, \infty \right]$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x_0 < x$ we have $\frac{\delta}{2} < f(x) \leqslant \infty \wedge \frac{\delta}{2} < g(x) \leqslant \infty \Rightarrow \delta < f(x) + g(x) \Rightarrow (f + g)(x) = f(x) + g(x) \in]\delta, \infty] \subseteq U$ which proves that $\lim_{x \downarrow x_0} (f + g)(x) = L_1 + L_2 = \lim_{x \downarrow x_0} f(x) + \lim_{x \downarrow x_0} g(x)$.

$L_1 \in \mathbb{R}, L_2 = \infty$. Let U be a open set with $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in]\delta, \infty] \subseteq U$ then as $L_2 = \infty$ we have $L_2 \in]2 \cdot \delta - L_1, \infty]$ and as $L_1 \in \mathbb{R}$ we have that $L_1 \in]L_1 - \delta, L_1 + \delta[$, so by the definition of a right limit we have that x_0 is the right limit point of A and that there exists open sets V_f, V_g containing x_0 such that $\forall x \in V_f \bigcap A$ with $x_0 < x$ we have $f(x) \in]L_1 - \delta, L_1 + \delta[$ and $\forall x \in V_g \bigcap A$ with $x_0 < x$ we have $g(x) \in]2 \cdot \delta - L_1, \infty]$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x_0 < x$ we have $L_1 - \delta + 2 \cdot \delta - L_1 < f(x) + g(x) \leqslant \infty \Rightarrow (f+g)(x) = f(x) + g(x) \in]\delta, \infty] \subseteq U$ which proves that $\lim_{x \downarrow x_0} (f+g)(x) = L_1 + L_2 = \lim_{x \downarrow x_0} f(x) + \lim_{x \downarrow x_0} g(x)$

$L_1 = \infty, L_2 \in \mathbb{R}$. Let U be a open set with $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in]\delta, \infty] \subseteq U$ then as $L_1 = \infty$ we have $L_1 \in]2 \cdot \delta - L_2, \infty]$ and as $L_2 \in \mathbb{R}$ we have that $L_2 \in]L_2 - \delta, L_2 + \delta[$, so by the definition of a right limit we have that x_0 is the right limit point of A and that there exists open sets V_f, V_g containing x_0 such that $\forall x \in V_g \bigcap A$ with $x_0 < x$ we have $g(x) \in]L_2 - \delta, L_2 + \delta[$ and $\forall x \in V_f \bigcap A$ with $x_0 < x$ we have $f(x) \in]2 \cdot \delta - L_2, \infty]$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x_0 < x$ we have $L_2 - \delta + 2 \cdot \delta - L_2 < f(x) + g(x) \leqslant \infty \Rightarrow (f+g)(x) = f(x) + g(x) \in]\delta, \infty] \subseteq U$ which proves that $\lim_{x \downarrow x_0} (f+g)(x) = L_1 + L_2 = \lim_{x \downarrow x_0} f(x) + \lim_{x \downarrow x_0} g(x)$

$L_1 + L_2 = -\infty$. then we have either

$L_1, L_2 = -\infty$. Let U be a open set with $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in [-\infty, -\delta[\subseteq U$ then as $L_1, L_2 = -\infty$ we have $L_1, L_2 \in \left[-\infty, -\frac{\delta}{2}\right[$ and by the definition of right limits we have that x_0 is a right limit point of A and there exists open sets V_f, V_g containing x_0 such that $\forall x \in V_f \bigcap A$ with $x_0 < x$ we have $f(x) \in \left[-\infty, -\frac{\delta}{2}\right[$ and $\forall x \in V_g \bigcap A$ with $x_0 < x$ we have $g(x) \in \left[-\infty, -\frac{\delta}{2}\right[$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x_0 < x$ we have $-\infty \leqslant f(x) < -\frac{\delta}{2} \wedge -\infty \leqslant g(x) < -\frac{\delta}{2} \Rightarrow -\infty \leqslant f(x) + g(x) \Rightarrow (f+g)(x) = f(x) + g(x) \in [-\infty, \delta[\subseteq U$ which proves that $\lim_{x \downarrow x_0} (f+g)(x) = L_1 + L_2 = \lim_{x \downarrow x_0} f(x) + \lim_{x \downarrow x_0} g(x)$.

$L_1 \in \mathbb{R}, L_2 = -\infty$. Let U be a open set with $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in [-\infty, -\delta[\subseteq U$ then as $L_2 = -\infty$ we have $L_2 \in [-\infty, -2 \cdot \delta - L_1[$ and as $L_1 \in \mathbb{R}$ we have that $L_1 \in]L_1 - \delta, L_1 + \delta[$, so by the definition of a right limit we have that x_0 is the right limit point of A and that there exists open sets V_f, V_g containing x_0 such that $\forall x \in V_f \bigcap A$

with $x_0 < x$ we have $f(x) \in]L_1 - \delta, L_1 + \delta[$ and $\forall x \in V_g \bigcap A$ with $x_0 < x$ we have $g(x) \in [-\infty, -2 \cdot \delta - L_1[$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x_0 < x$ we have $-\infty \leqslant f(x) + g(x) < L_1 + \delta - 2 \cdot \delta - L_1 \Rightarrow (f + g)(x) = f(x) + g(x) \in [-\infty, -\delta[\subseteq U$ which proves that $\lim_{x \downarrow x_0} (f + g)(x) = L_1 + L_2 = \lim_{x \downarrow x_0} f(x) + \lim_{x \downarrow x_0} g(x)$

$\boldsymbol{L_1 = -\infty, L_2 \in \mathbb{R}.}$ Let U be a open set with $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in [-\infty, -\delta[\subseteq U$ then as $L_1 = \infty$ we have $L_1 \in [-\infty, -2 \cdot \delta - L_2[$ and as $L_2 \in \mathbb{R}$ we have that $L_2 \in]L_2 - \delta, L_2 + \delta[$, so by the definition of a right limit we have that x_0 is the right limit point of A and that there exists open sets V_f, V_g containing x_0 such that $\forall x \in V_g \bigcap A$ with $x_0 < x$ we have $g(x) \in]L_2 - \delta, L_2 + \delta[$ and $\forall x \in V_f \bigcap A$ with $x_0 < x$ we have $f(x) \in [-\infty, -2 \cdot \delta - L_2[$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x_0 < x$ we have $-\infty \leqslant f(x) + g(x) < -2 \cdot \delta - L_2 + L_2 + \delta \Rightarrow (f + g)(x) = f(x) + g(x) \in [-\infty, -\delta[\subseteq U$ which proves that $\lim_{x \downarrow x_0} (f + g)(x) = L_1 + L_2 = \lim_{x \downarrow x_0} f(x) + \lim_{x \downarrow x_0} g(x)$

5. Let U be a open set containing $L_1 + L_2$ then we have the following cases to consider for $L_1 + L_2$

$\boldsymbol{L_1 + L_2 \in \mathbb{R}.}$ then we must have $L_1, L_2 \in \mathbb{R}$, If now $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in]L_1 + L_2 - \varepsilon, L_1 + L_2 - \varepsilon[\subseteq U$. From $L_1 \in \left]L_1 - \frac{\varepsilon}{2}, L_1 + \frac{\varepsilon}{2}\right[$, $L_2 \in \left]L_2 - \frac{\varepsilon}{2}, L_2 + \frac{\varepsilon}{2}\right[$ and the definition of left limits we have that x_0 is a left limit point of A and there exists open sets V_f, V_g with $x_0 \in V_f, V_g$ such that $\forall x \in V_f \bigcap A$ with $x < x_0$ we have $f(x) \in \left]L_1 - \frac{\varepsilon}{2}, L_2 - \frac{\varepsilon}{2}\right[$ and $\forall x \in V_g \bigcap A$ with $x < x_0$ we have $g(x) \in \left]L_2 - \frac{\varepsilon}{2}, L_2 + \frac{\varepsilon}{2}\right[$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x < x_0$ then $L_1 - \frac{\varepsilon}{2} < f(x) < L_1 + \frac{\varepsilon}{2} \wedge L_2 - \frac{\varepsilon}{2} < g(x) < L_2 + \frac{\varepsilon}{2} \Rightarrow L_1 + L_2 - \varepsilon < f(x) + g(x) < L_1 + L_2 + \varepsilon \Rightarrow (f + g)(x) = f(x) + g(x) \in]L_1 + L_2 + \varepsilon, L_1 + L_2 - \varepsilon[\subseteq U$ which proves that $\lim_{x \uparrow x_0} (f + g)(x) = L_1 + L_2 = \lim_{x \uparrow x_0} f(x) + \lim_{x \uparrow x_0} g(x)$.

$\boldsymbol{L_1 + L_2 = \infty.}$ then we have either

$\boldsymbol{L_1, L_2 = \infty.}$ Let U be a open set with $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in]\delta, \infty] \subseteq U$ then as $L_1, L_2 = \infty$ we have $L_1, L_2 \in \left]\frac{\delta}{2}, \infty\right]$ and by the definition of left limits we have that x_0 is a left limit point of A and there exists open sets V_f,V_g containing x_0 such that $\forall x \in V_f \bigcap A$ with $x < x_0$ we have $f(x) \in \left]\frac{\delta}{2}, \infty\right]$ and $\forall x \in V_g \bigcap A$ with $x < x_0$ we have $g(x) \in \left]\frac{\delta}{2}, \infty\right]$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x < x_0$ we have $\frac{\delta}{2} < f(x) \leqslant \infty \wedge \frac{\delta}{2} < g(x) \leqslant \infty \Rightarrow \delta < f(x) + g(x) \Rightarrow (f + g)(x) = f(x) + g(x) \in]\delta, \infty] \subseteq U$ which proves that $\lim_{x \uparrow x_0} (f + g)(x) = L_1 + L_2 = \lim_{x \uparrow x_0} f(x) + \lim_{x \uparrow x_0} g(x)$.

$L_1 \in \mathbb{R}, L_2 = \infty$. Let U be a open set with $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in]\delta, \infty] \subseteq U$ then as $L_2 = \infty$ we have $L_2 \in]2 \cdot \delta - L_1, \infty]$ and as $L_1 \in \mathbb{R}$ we have that $L_1 \in]L_1 - \delta, L_1 + \delta[$, so by the definition of a left limit we have that x_0 is the left limit point of A and that there exists open sets V_f, V_g containing x_0 such that $\forall x \in V_f \bigcap A$ with $x < x_0$ we have $f(x) \in]L_1 - \delta, L_1 + \delta[$ and $\forall x \in V_g \bigcap A$ with $x < x_0$ we have $g(x) \in]2 \cdot \delta - L_1, \infty]$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x < x_0$ we have $L_1 - \delta + 2 \cdot \delta - L_1 < f(x) + g(x) \leqslant \infty \Rightarrow (f+g)(x) = f(x) + g(x) \in]\delta, \infty] \subseteq U$ which proves that $\lim_{x \uparrow x_0} (f+g)(x) = L_1 + L_2 = \lim_{x \uparrow x_0} f(x) + \lim_{x \uparrow x_0} g(x)$

$L_1 = \infty, L_2 \in \mathbb{R}$. Let U be a open set with $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in]\delta, \infty] \subseteq U$ then as $L_1 = \infty$ we have $L_1 \in]2 \cdot \delta - L_2, \infty]$ and as $L_2 \in \mathbb{R}$ we have that $L_2 \in]L_2 - \delta, L_2 + \delta[$, so by the definition of a left limit we have that x_0 is the left limit point of A and that there exists open sets V_f, V_g containing x_0 such that $\forall x \in V_g \bigcap A$ with $x < x_0$ we have $g(x) \in]L_2 - \delta, L_2 + \delta[$ and $\forall x \in V_f \bigcap A$ with $x < x_0$ we have $f(x) \in]2 \cdot \delta - L_2, \infty]$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x < x_0$ we have $L_2 - \delta + 2 \cdot \delta - L_2 < f(x) + g(x) \leqslant \infty \Rightarrow (f+g)(x) = f(x) + g(x) \in]\delta, \infty] \subseteq U$ which proves that $\lim_{x \uparrow x_0} (f+g)(x) = L_1 + L_2 = \lim_{x \uparrow x_0} f(x) + \lim_{x \uparrow x_0} g(x)$

$L_1 + L_2 = -\infty$. then we have either

$L_1, L_2 = -\infty$. Let U be a open set with $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in [-\infty, -\delta[\subseteq U$ then as $L_1, L_2 = -\infty$ we have $L_1, L_2 \in \left[-\infty, -\frac{\delta}{2}\right[$ and by the definition of left limits we have that x_0 is a left limit point of A and there exists open sets V_f, V_g containing x_0 such that $\forall x \in V_f \bigcap A$ with $x < x_0$ we have $f(x) \in \left[-\infty, -\frac{\delta}{2}\right[$ and $\forall x \in V_g \bigcap A$ with $x < x_0$ we have $g(x) \in \left[-\infty, -\frac{\delta}{2}\right[$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x < x_0$ we have $-\infty \leqslant f(x) < -\frac{\delta}{2} \wedge -\infty \leqslant g(x) < -\frac{\delta}{2} \Rightarrow -\infty \leqslant f(x) + g(x) \Rightarrow (f+g)(x) = f(x) + g(x) \in [-\infty, \delta[\subseteq U$ which proves that $\lim_{x \uparrow x_0} (f+g)(x) = L_1 + L_2 = \lim_{x \uparrow x_0} f(x) + \lim_{x \uparrow x_0} g(x)$.

$L_1 \in \mathbb{R}, L_2 = -\infty$. Let U be a open set with $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in [-\infty, -\delta[\subseteq U$ then as $L_2 = -\infty$ we have $L_2 \in [-\infty, -2 \cdot \delta - L_1[$ and as $L_1 \in \mathbb{R}$ we have that $L_1 \in]L_1 - \delta, L_1 + \delta[$, so by the definition of a left limit we have that x_0 is the left limit point of A and that there exists open sets V_f, V_g containing x_0 such that $\forall x \in V_f \bigcap A$ with $x < x_0$ we have $f(x) \in]L_1 - \delta, L_1 + \delta[$ and $\forall x \in V_g \bigcap A$ with $x < x_0$ we have $g(x) \in [-\infty, -2 \cdot \delta - L_1[$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x < x_0$ we have $-\infty \leqslant f(x) + g(x) < L_1 + \delta - 2 \cdot \delta - L_1 \Rightarrow$

$(f+g)(x) = f(x) + g(x) \in [-\infty, -\delta[\subseteq U$ which proves that $\lim_{x \uparrow x_0} (f+g)(x) = L_1 + L_2 = \lim_{x \uparrow x_0} f(x) + \lim_{x \uparrow x_0} g(x)$

$\boldsymbol{L_1 = -\infty, L_2 \in \mathbb{R}.}$ Let U be a open set with $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in [-\infty, -\delta[\subseteq U$ then as $L_1 = \infty$ we have $L_1 \in [-\infty, -2 \cdot \delta - L_2[$ and as $L_2 \in \mathbb{R}$ we have that $L_2 \in]L_2 - \delta, L_2 + \delta[$, so by the definition of a left limit we have that x_0 is the left limit point of A and that there exists open sets V_f, V_g containing x_0 such that $\forall x \in V_g \bigcap A$ with $x < x_0$ we have $g(x) \in]L_2 - \delta, L_2 + \delta[$ and $\forall x \in V_f \bigcap A$ with $x < x_0$ we have $f(x) \in [-\infty, -2 \cdot \delta - L_2[$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x < x_0$ we have $-\infty \leqslant f(x) + g(x) < -2 \cdot \delta - L_2 + L_2 + \delta \Rightarrow (f+g)(x) = f(x) + g(x) \in [-\infty, -\delta[\subseteq U$ which proves that $\lim_{x \uparrow x_0} (f+g)(x) = L_1 + L_2 = \lim_{x \uparrow x_0} f(x) + \lim_{x \uparrow x_0} g(x)$

6. Let U be a open set containing $L_1 + L_2$ then we have the following cases to consider for $L_1 + L_2$

$\boldsymbol{L_1 + L_2 \in \mathbb{R}.}$ then we must have $L_1, L_2 \in \mathbb{R}$, If now $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in]L_1 + L_2 - \varepsilon, L_1 + L_2 - \varepsilon[\subseteq U$. From $L_1 \in \left]L_1 - \frac{\varepsilon}{2}, L_1 + \frac{\varepsilon}{2}\right[$, $L_2 \in \left]L_2 - \frac{\varepsilon}{2}, L_2 + \frac{\varepsilon}{2}\right[$ and the definition of limits we have that x_0 is a limit point of A and there exists open sets V_f, V_g with $x_0 \in V_f, V_g$ such that $\forall x \in V_f \bigcap A$ with $x \neq x_0$ we have $f(x) \in \left]L_1 - \frac{\varepsilon}{2}, L_2 - \frac{\varepsilon}{2}\right[$ and $\forall x \in V_g \bigcap A$ with $x \neq x_0$ we have $g(x) \in \left]L_2 - \frac{\varepsilon}{2}, L_2 + \frac{\varepsilon}{2}\right[$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x \neq x_0$ then $L_1 - \frac{\varepsilon}{2} < f(x) < L_1 + \frac{\varepsilon}{2} \wedge L_2 - \frac{\varepsilon}{2} < g(x) < L_2 + \frac{\varepsilon}{2} \Rightarrow L_1 + L_2 - \varepsilon < f(x) + g(x) < L_1 + L_2 + \varepsilon \Rightarrow (f+g)(x) = f(x) + g(x) \in]L_1 + L_2 + \varepsilon, L_1 + L_2 - \varepsilon[\subseteq U$ which proves that $\lim_{x \rightarrow x_0} (f+g)(x) = L_1 + L_2 = \lim_{x \rightarrow x_0} f(x) + \lim_{x \rightarrow x_0} g(x)$.

$\boldsymbol{L_1 + L_2 = \infty.}$ then we have either

$\boldsymbol{L_1, L_2 = \infty.}$ Let U be a open set with $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in]\delta, \infty] \subseteq U$ then as $L_1, L_2 = \infty$ we have $L_1, L_2 \in \left]\frac{\delta}{2}, \infty\right]$ and by the definition of limits we have that x_0 is a limit point of A and there exists open sets V_f, V_g containing x_0 such that $\forall x \in V_f \bigcap A$ with $x \neq x_0$ we have $f(x) \in \left]\frac{\delta}{2}, \infty\right]$ and $\forall x \in V_g \bigcap A$ with $x \neq x_0$ we have $g(x) \in \left]\frac{\delta}{2}, \infty\right]$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x \neq x_0$ we have $\frac{\delta}{2} < f(x) \leqslant \infty \wedge \frac{\delta}{2} < g(x) \leqslant \infty \Rightarrow \delta < f(x) + g(x) \Rightarrow (f+g)(x) = f(x) + g(x) \in]\delta, \infty] \subseteq U$ which proves that $\lim_{x \rightarrow x_0} (f+g)(x) = L_1 + L_2 = \lim_{x \rightarrow x_0} f(x) + \lim_{x \rightarrow x_0} g(x)$.

$\boldsymbol{L_1 \in \mathbb{R}, L_2 = \infty.}$ Let U be a open set with $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in]\delta, \infty] \subseteq U$ then as $L_2 = \infty$ we have $L_2 \in]2 \cdot \delta - L_1, \infty]$ and as $L_1 \in \mathbb{R}$ we have that $L_1 \in]L_1 - \delta, L_1 + \delta[$, so by the definition of a limit we have that x_0 is the limit point of A and that there exists open sets V_f, V_g containing x_0 such that $\forall x \in V_f \bigcap A$ with $x \neq x_0$ we have $f(x) \in]L_1 - \delta, L_1 + \delta[$ and $\forall x \in V_g \bigcap A$ with $x \neq x_0$ we have $g(x) \in]2 \cdot \delta - L_1, \infty]$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x_0 \neq x$ we have $L_1 - \delta + 2 \cdot \delta - L_1 < f(x) + g(x) \leqslant \infty \Rightarrow (f+g)(x) = f(x) + g(x) \in]\delta, \infty] \subseteq U$ which proves that $\lim_{x \to x_0} (f+g)(x) = L_1 + L_2 = \lim_{x \to x_0} f(x) + \lim_{x \to x_0} g(x)$

$L_1 = \infty, L_2 \in \mathbb{R}$. Let U be a open set with $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in]\delta, \infty] \subseteq U$ then as $L_1 = \infty$ we have $L_1 \in]2 \cdot \delta - L_2, \infty]$ and as $L_2 \in \mathbb{R}$ we have that $L_2 \in]L_2 - \delta, L_2 + \delta[$, so by the definition of a limit we have that x_0 is the limit point of A and that there exists open sets V_f, V_g containing x_0 such that $\forall x \in V_g \bigcap A$ with $x \neq x_0$ we have $g(x) \in]L_2 - \delta, L_2 + \delta[$ and $\forall x \in V_f \bigcap A$ with $x \neq x_0$ we have $f(x) \in]2 \cdot \delta - L_2, \infty]$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x \neq x_0$ we have $L_2 - \delta + 2 \cdot \delta - L_2 < f(x) + g(x) \leqslant \infty \Rightarrow (f+g)(x) = f(x) + g(x) \in]\delta, \infty] \subseteq U$ which proves that $\lim_{x \to x_0} (f+g)(x) = L_1 + L_2 = \lim_{x \to x_0} f(x) + \lim_{x \to x_0} g(x)$

$L_1 + L_2 = -\infty$. then we have either

$L_1, L_2 = -\infty$. Let U be a open set with $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in [-\infty, -\delta[\subseteq U$ then as $L_1, L_2 = -\infty$ we have $L_1, L_2 \in \left[-\infty, -\frac{\delta}{2}\right[$ and by the definition of limits we have that x_0 is a limit point of A and there exists open sets V_f, V_g containing x_0 such that $\forall x \in V_f \bigcap A$ with $x_0 < x$ we have $f(x) \in \left[-\infty, -\frac{\delta}{2}\right[$ and $\forall x \in V_g \bigcap A$ with $x \neq x$ we have $g(x) \in \left[-\infty, -\frac{\delta}{2}\right[$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x \neq x_0$ we have $-\infty \leqslant f(x) < -\frac{\delta}{2} \wedge -\infty \leqslant g(x) < -\frac{\delta}{2} \Rightarrow -\infty \leqslant f(x) + g(x) \Rightarrow (f+g)(x) = f(x) + g(x) \in [-\infty, \delta[\subseteq U$ which proves that $\lim_{x \to x_0} (f+g)(x) = L_1 + L_2 = \lim_{x \to x_0} f(x) + \lim_{x \to x_0} g(x)$.

$L_1 \in \mathbb{R}, L_2 = -\infty$. Let U be a open set with $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in [-\infty, -\delta[\subseteq U$ then as $L_2 = -\infty$ we have $L_2 \in [-\infty, -2 \cdot \delta - L_1[$ and as $L_1 \in \mathbb{R}$ we have that $L_1 \in]L_1 - \delta, L_1 + \delta[$, so by the definition of a limit we have that x_0 is the limit point of A and that there exists open sets V_f, V_g containing x_0 such that $\forall x \in V_f \bigcap A$ with $x \neq x_0$ we have $f(x) \in]L_1 - \delta, L_1 + \delta[$ and $\forall x \in V_g \bigcap A$ with $x \neq x$ we have $g(x) \in [-\infty, -2 \cdot \delta - L_1[$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x \neq x_0$ we have $-\infty \leqslant f(x) + g(x) < L_1 + \delta - 2 \cdot \delta - L_1 \Rightarrow (f+g)(x) = f(x) + g(x) \in [-\infty, -\delta[\subseteq U$ which proves that $\lim_{x \to x_0} (f+g)(x) = L_1 + L_2 = \lim_{x \to x_0} f(x) + \lim_{x \to x_0} g(x)$

$L_1 = -\infty, L_2 \in \mathbb{R}$. Let U be a open set with $L_1 + L_2 \in U$ a open set then there exists a $\delta > 0$ such that $L_1 + L_2 \in [-\infty, -\delta[\subseteq U$ then as $L_1 = \infty$ we have $L_1 \in [-\infty, -2 \cdot \delta - L_2[$ and as $L_2 \in \mathbb{R}$ we have that $L_2 \in]L_2 - \delta, L_2 + \delta[$, so by the definition of a limit we have that x_0 is the limit point of A and that there exists open sets V_f, V_g containing x_0 such that $\forall x \in V_g \bigcap A$ with $x \neq x_0$ we have $g(x) \in]L_2 - \delta, L_2 + \delta[$ and $\forall x \in V_f \bigcap A$ with $x \neq x_0$ we have $f(x) \in [-\infty, -2 \cdot \delta - L_2[$. So if $x \in (V_f \bigcap V_g) \bigcap A$ with $x = x_0$ we have $-\infty \leqslant f(x) + g(x) < -2 \cdot \delta - L_2 + L_2 + \delta \Rightarrow (f+g)(x) = f(x) + g(x) \in [-\infty, -\delta[\subseteq U$ which proves that $\lim_{x \to x_0} (f+g)(x) = L_1 + L_2 = \lim_{x \to x_0} f(x) + \lim_{x \to x_0} g(x)$. □

Theorem 17.52. *Let $A, B \subseteq \bar{\mathbb{R}}$, $f: A \to B$ a function $x_0, L \in \bar{\mathbb{R}}$ then $\lim_{x \downarrow x_0} f(x) = L = \lim_{x \uparrow x_0} f(x)$ if and only if $\lim_{x \to x_0} f(x) = L$ and x_0 is a left and right limit point of A.*

Proof.

⇒. From the fact that the left and right limits exists we have that x_0 is a left and right limit point of A. So if U open set contains x_0 there exists a $x \in U \bigcap A$ with $x < x_0 \Rightarrow x \neq x_0$ proving that x_0 is a limit point of A. Take $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ then there exists a $V_1, V_2 \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $x_0 \in V_1, V_2$ such that $\forall x \in V_1 \bigcap A \vdash x < x_0$ we have $f(x) \in U$ and $\forall x \in V_2 \bigcap A \vdash x_0 < x$ we have $f(x) \in U$. So if we take $W = V_1 \bigcap V_2 \in \mathcal{T}_{\bar{\mathbb{R}}}$ then $x_0 \in W$ and $\forall x \in W \bigcap A \vdash x \neq x_0$ we have $x < x_0 \vee x_0 < x$ giving as x_0 is proven to be a limit point that $\lim_{x \to x_0} f(x) = L$.

⇐. Take $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ then there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $x_0 \in V$ such that $\forall x \in V \bigcap A \vdash x \neq x_0$ we have $f(x) \in U$ proving that $\forall x \in V \bigcap A \vdash x < x_0$ [or $x_0 < x$] $\Rightarrow x \neq x_0$ giving $f(x) \in U$. which proves, as we assume that x_0 is a left and right limit point of A, that $\lim_{x \downarrow x_0} f(x) = L = \lim_{x \uparrow x_0} f(x)$. □

Note 17.53. The extra condition on the right side of the above equivalence is needed because the fact that x_0 is a limit point of A does not guarantees that x_0 is a left and right limit point of A. For example in $A =]1, \infty]$ we have that ∞ is a limit point of A, also a left limit point of A but not a right limit point of A. So a well defined right limit at ∞ can not exists. However we can have at ∞ and $-\infty$ the following equivalences.

The following theorem essentially proves that limits preserver order on functions.

Theorem 17.54. *Let $A, B \subseteq \bar{\mathbb{R}}$, $f: A \to B$ a function, $L \in \bar{\mathbb{R}}$ then we have*

1. *$\lim_{x \to -\infty} f(x) = L$ if and only if $\lim_{x \downarrow -\infty} f = L$*
2. *$\lim_{x \to \infty} f(x) = L$ if and onlyy if $\lim_{x \uparrow \infty} f = L$*

Proof.

1.

   ⇒. As $-\infty$ is a limit point of A we have that $\forall U$ open with $-\infty \in U$ there exists a $x \in U \bigcap A$ with $x \neq -\infty \Rightarrow -\infty < x$ proving that $-\infty$ is a right limit point of A. Further if U is a open set containing L then we have that there exists a open set V containing $-\infty$ such that $\forall x \in V \bigcap A$ with $x \neq -\infty$ we have $f(x) \in U$, as $x \neq -\infty$ implies $-\infty < x$ we have then proved that $\lim_{x \downarrow -\infty} f(x) = L$.

   ⇐. As $-\infty$ is a right limit point of A we have $\forall U$ open containing $-\infty$ there exists a $x \in U \bigcap A$ with $-\infty < x \Rightarrow -\infty \neq x$ proving that $-\infty$ is a limit point of A. Further if U is a open set containing L then there exists a open set V containing $-\infty$ such that $\forall x \in V \bigcap A$ with $-\infty < x$ we have $f(x) \in U$, as $x < -\infty$ implies $x \neq -\infty$ we have $\lim_{x \rightarrow -\infty} f(x)$

2.

   ⇒. As ∞ is a limit point of A we have that $\forall U$ open with $\infty \in U$ there exists a $x \in U \bigcap A$ with $x \neq \infty \Rightarrow x < \infty$ proving that ∞ is a left limit point of A. Further if U is a open set containing L then we have that there exists a open set V containing ∞ such that $\forall x \in V \bigcap A$ with $x \neq \infty$ we have $f(x) \in U$, as $x \neq \infty$ implies $x < \infty$ we have then proved that $\lim_{x \uparrow \infty} f(x) = L$.

   ⇐. As ∞ is a left limit point of A we have $\forall U$ open containing ∞ there exists a $x \in U \bigcap A$ with $x < \infty \Rightarrow \infty \neq x$ proving that ∞ is a limit point of A. Further if U is a open set containing L then there exists a open set V containing ∞ such that $\forall x \in V \bigcap A$ with $x < \infty$ we have $f(x) \in U$, as $x < \infty$ implies $x \neq \infty$ we have $\lim_{x \rightarrow \infty} f(x)$ □

Note that the limit of a function is related to continuity at a point in the topology $\mathcal{T}_{\bar{\mathbb{R}}}$ as the following theorem shows.

Theorem 17.55. *Let $A, B \subseteq \bar{\mathbb{R}}$, $f: A \rightarrow B$ a function $x_0 \in A$ a limit point of x_0 then f is continuous at x_0 (see 12.131) if and only if $\lim_{x \rightarrow x_0} f(x) = f(x_0)$*

Proof. Let $\mathcal{T}_A = \{U \bigcap A | U \in \mathcal{T}_{\bar{\mathbb{R}}}\}$ be the subspace topology of A, and $\mathcal{T}_B = \{U \bigcap B | U \in \mathcal{T}_{\bar{\mathbb{R}}}\}$ be the subspace topology of B.

⇒. Take now $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $f(x_0) \in U \underset{f(x_0) \in B}{\Rightarrow} f(x_0) \in U \bigcap B \in \mathcal{T}_B$, so by continuity of f in the subspace topologies we have $\exists V \in \mathcal{T}_A$ such that $x_0 \in V \wedge f(V) \subseteq U \bigcap B$. As $V \in \mathcal{T}_A$ there exists a $V' \in \mathcal{T}_{\bar{\mathbb{R}}}$ such that $V = V' \bigcap A$. So if $x \in A \bigcap V' \vdash x \neq x_0$ we have $x \in V \Rightarrow f(x) \in U \bigcap B \subseteq U$ proving that $\lim_{x \rightarrow x_0} f(x) = f(x_0)$

⇐. Take $U \in \mathcal{T}_B$ with $f(x_0) \in U$ then $\exists U' \in \mathcal{T}_{\bar{\mathbb{R}}}$ such that $U = U' \bigcap B$ giving $f(x_0) \in U'$ so there exists a $V' \in \mathcal{T}_{\bar{\mathbb{R}}}$ such that $x_0 \in V'$ and $\forall x \in V' \bigcap A \vdash x \neq x_0$ we have $f(x) \in U' \Rightarrow f(x) \in U$ [as $f(x) \in B$]. Take now $V = V' \bigcap A \in \mathcal{T}_A$ then as $x_0 \in A$ we have $x_0 \in V$ and $f(x_0) \in U$ [as $f(x_0) \in B$, also if $x \in V \vdash x \neq x_0$ then $f(x) \in U$ so that $f(V) \subseteq U$ proving that f is continuous a x_0. □

The above theorem motivates the definition of left or right continuity

Definition 17.56. *Let $A, B \subseteq \bar{\mathbb{R}}$, $f: A \rightarrow B$, $x_0 \in A$ a left [right] limit point of A then f is left [right] continuous at x_0 iff $\lim_{x \uparrow x_0} f(x) = f(x_0)$ [$\lim_{x \downarrow x_0} f(x) = f(x_0)$]*

We now look at some special cases of limits where x_0 or L where in $\mathbb{R}$ or equal to ∞, $-\infty$

Theorem 17.57. *:et $A, B \subseteq \bar{\mathbb{R}}$, $f: A \rightarrow B$ and $x_0, L \in \bar{\mathbb{R}}$ then we have the following possible cases for a left, right, limit*

$x_0 \in \mathbb{R}, L \in \mathbb{R}$. *then we have for the different limits the following alternative definitions*

$\lim_{x \uparrow x_0} f(x) = L$. *$x_0$ is a left limit point of A and $\forall \varepsilon > 0 \vDash \exists \delta > 0$ such that $\forall x \in A \vdash 0 < x_0 - x < \delta$ we have $|f(x) - L| < \varepsilon$*

$\lim_{x \downarrow x_0} f(x) = L$. *$x_0$ is a right limit point of A and $\forall \varepsilon > 0 \vDash \exists \delta > 0$ such that $\forall x \in A \vdash 0 < x - x_0 < \delta$ we have $|f(x) - L| < \varepsilon$*

$\lim_{x \rightarrow x_0} f(x) = L$. *$x_0$ is a limit point of A and $\forall \varepsilon > 0 \vDash \exists \delta > 0$ such that $\forall x \in A \vdash 0 < |x_0 - x| < \delta$ we have $|f(x) - L| < \varepsilon$*

$x_0 \in \mathbb{R}, L = \infty$. *then we have for the different limits the following alternative definitions*

$\lim_{x \uparrow x_0} f(x) = \infty$. *$x_0$ is a left limit point of A and $\forall \varepsilon > 0$ there exists a $\delta > 0$ such that if $x \in A \wedge 0 < x_0 - x < \delta$ we have $\varepsilon < f(x)$*

$\lim_{x \downarrow x_0} f(x) = \infty$. *$x_0$ is a right limit point of A and $\forall \varepsilon > 0$ there exists a $\delta > 0$ such that if $x \in A \wedge 0 < x - x_0 < \delta$ we have $\varepsilon < f(x)$*

$\lim_{x \rightarrow x_0} f(x) = \infty$. *$x_0$ is a limit point of A and $\forall \varepsilon > 0$ there exists a $\delta > 0$ such that if $x \in A \wedge 0 < |x_0 - x| < \delta$ we have $\varepsilon < f(x)$*

$x_0 \in \mathbb{R}, L = -\infty$. *then we have for the different limits the following alternative definitions*

$\lim_{x \uparrow x_0} f(x) = -\infty$. *$x_0$ is a left limit point of A and $\forall \varepsilon > 0$ there exists a $\delta > 0$ such that if $x \in A \wedge 0 < x_0 - x < \delta$ we have $f(x) < -\varepsilon$*

$\lim_{x \downarrow x_0} f(x) = -\infty$. *$x_0$ is a right limit point of A and $\forall \varepsilon > 0$ there exists a $\delta > 0$ such that if $x \in A \wedge 0 < x - x_0 < \delta$ we have $f(x) < -\varepsilon$*

$\lim_{x \rightarrow x_0} f(x) = -\infty$. *$x_0$ is a limit point of A and $\forall \varepsilon > 0$ there exists a $\delta > 0$ such that if $x \in A \wedge 0 < |x_0 - x| < \delta$ we have $f(x) < -\varepsilon$*

$x_0 = \infty, L \in \mathbb{R}$. *then we have for the different limits the following alternative definitions*

$\lim_{x \uparrow \infty} f(x) = L$. *$\infty$ is a left limit point of A and $\forall \varepsilon > 0$ there exists a $\delta > 0$ such that if $x \in A \wedge \delta < x < \infty$ we have $|f(x) - L| < \varepsilon$*

$\boldsymbol{\lim_{x\downarrow\infty} f(x) = L}$. *as ∞ can not be a right limit point of any set A this can not happen.*

$\boldsymbol{\lim_{x\rightarrow\infty} f(x) = L}$. $\lim_{x\uparrow\infty} f(x) = L$

$\boldsymbol{x_0 = \infty, L = \infty}$. *then we have for the different limits the following alternative definitions*

$\boldsymbol{\lim_{x\uparrow\infty} f(x) = \infty}$. *$\infty$ is a left limit point of A and $\forall\varepsilon > 0$ there exists a $\delta > 0$ so that if $\delta < x < \infty$ then $\varepsilon < f(x)$*

$\boldsymbol{\lim_{x\downarrow\infty} f(x) = \infty}$. *as ∞ can not be a right limit point of any set A this can not happen.*

$\boldsymbol{\lim_{x\rightarrow\infty} f(x) = \infty}$. $\lim_{x\uparrow\infty} f(x) = \infty$

$\boldsymbol{x_0 = \infty, L = -\infty}$. *then we have for the different limits the following alternative definitions*

$\boldsymbol{\lim_{x\uparrow\infty} f(x) = -\infty}$. *$\infty$ is a left limit point of A and $\forall\varepsilon > 0$ there exists a $\delta > 0$ so that if $\delta < x < \infty$ then $f(x) < -\varepsilon$*

$\boldsymbol{\lim_{x\downarrow\infty} f(x) = -\infty}$. *as ∞ can not be a right limit point of any set A this can not happen.*

$\boldsymbol{\lim_{x\rightarrow\infty} f(x) = -\infty}$. $\lim_{x\uparrow\infty} f(x) = -\infty$

$\boldsymbol{x_0 = -\infty, L \in \mathbb{R}}$. *then we have for the different limits the following alternative definitions*

$\boldsymbol{\lim_{x\uparrow-\infty} f(x) = L}$. *as $-\infty$ can not be a left limit point of any set A this can not happen.*

$\boldsymbol{\lim_{x\downarrow-\infty} f(x) = L}$. *$-\infty$ is a right limit point of A and $\forall\varepsilon > 0$ there exists a $\delta > 0$ so that if $-\infty < x < -\infty$ then $|f(x) - L| < \varepsilon$*

$\boldsymbol{\lim_{x\rightarrow-\infty} f(x) = L}$. $\lim_{x\uparrow\infty} f(x) = L$

$\boldsymbol{x_0 = -\infty, L = \infty}$. *then we have for the different limits the following alternative definitions*

$\boldsymbol{\lim_{x\uparrow-\infty} f(x) = \infty}$. *as $-\infty$ can not be a left limit point of any set A this can not happen.*

$\boldsymbol{\lim_{x\downarrow-\infty} f(x) = \infty}$. *$-\infty$ is a right limit point of A and $\forall\varepsilon > 0$ there exists a $\delta > 0$ such that $\forall x \in A$ with $-\infty < x < -\delta$ we have $\varepsilon < f(x)$*

$\boldsymbol{\lim_{x\rightarrow-\infty} f(x) = \infty}$. $\lim_{x\downarrow-\infty} f(x) = \infty$

$\boldsymbol{x_0 = -\infty, L = -\infty}$. *then we have for the different limits the following alternative definitions*

$\boldsymbol{\lim_{x\uparrow-\infty} f(x) = -\infty}$. *as $-\infty$ can not be a left limit point of any set A this can not happen.*

$\lim\limits_{x\downarrow-\infty} f(x)=-\infty$**.** *$-\infty$ is a left limit point of A and $\forall\varepsilon>0$ there exists a $\delta>0$ such that $\forall x\in A$ with $-\infty<x<-\infty$ we have $f(x)<-\infty$*

$\lim\limits_{x\rightarrow-\infty} f(x)=-\infty$**.** $\lim\limits_{x\downarrow-\infty} f(x)=-\infty$

Proof.

$x_0\in\mathbb{R}, L\in\mathbb{R}$**.** then we have for the different limits the following alternative definitions

$\lim\limits_{x\uparrow\;x_0} f(x)=L$**.**

$\Rightarrow$**.** Given $\varepsilon>0$ take $L\in]L-\varepsilon,L+\varepsilon[\in\mathcal{T}_{\mathbb{R}}\subseteq\mathcal{T}_{\bar{\mathbb{R}}}$ [as $L\in\mathbb{R}$] then there exists a $V\in\mathcal{T}_{\bar{\mathbb{R}}}$ with $x_0\in V$ such that $\forall x\in V\bigcap A\vdash x<x_0$ we have $f(x)\in]L-\varepsilon,L+\varepsilon[\Rightarrow|f(x)-L|<\varepsilon$, as $x_0\in V\underset{x_0\in\mathbb{R}}{\Rightarrow}x_0\in V\bigcap\mathbb{R}\in\mathcal{T}_{\mathbb{R}}$ [using 17.36] $\Rightarrow\exists\delta>0\vdash x_0\in]x_0-\delta,x_0+\delta[\subseteq V\bigcap\mathbb{R}\subseteq V$. So if $x\in A\vdash 0<x_0-x<\delta$ then $x\in A\bigcap V\wedge x<x_0$ giving $|f(x)-L|<\varepsilon$.

$\Leftarrow$**.** Given $U\in\mathcal{T}_{\bar{\mathbb{R}}}\vdash L\in U$ then as $L\in\mathbb{R}$ we have $L\in U\bigcap\mathbb{R}\in\mathcal{T}_{\mathbb{R}}$ [using 17.36] so that $\exists\varepsilon>0$ such that $L\in]L-\varepsilon,L+\varepsilon[\subseteq U\bigcap\mathbb{R}$, by the hypothese there $\exists\delta>0$ such that if $x\in A\vdash 0<x_0-x<\delta$ we have $|f(x)-L|<\varepsilon\Rightarrow f(x)\in]L-\varepsilon,L+\varepsilon[\subseteq U\bigcap\mathbb{R}\subseteq U$. So if we take $V=]x_0-\delta,x_0+\delta[\in\mathcal{T}_{\mathbb{R}}\subseteq\mathcal{T}_{\bar{\mathbb{R}}}$ then $x_0\in V$ and if $x\in A\bigcap V$ with $x<x_0$ we have $0<x_0-x<\delta\wedge x\in A$, so that $f(x)\in U$. Proving that $\lim\limits_{x\uparrow x_0} f(x)=L$.

$\lim\limits_{x\downarrow x_0} f(x)=L$**.**

$\Rightarrow$**.** Given $\varepsilon>0$ take $L\in]L-\varepsilon,L+\varepsilon[\in\mathcal{T}_{\mathbb{R}}\subseteq\mathcal{T}_{\bar{\mathbb{R}}}$ [as $L\in\mathbb{R}$] then there exists a $V\in\mathcal{T}_{\bar{\mathbb{R}}}$ with $x_0\in V$ such that $\forall x\in V\bigcap A\vdash x_0<x$ we have $f(x)\in]L-\varepsilon,L+\varepsilon[\Rightarrow|f(x)-L|<\varepsilon$, as $x_0\in V\underset{x_0\in\mathbb{R}}{\Rightarrow}x_0\in V\bigcap\mathbb{R}\in\mathcal{T}_{\mathbb{R}}$ [using 17.36] $\Rightarrow\exists\delta>0\vdash x_0\in]x_0-\delta,x_0+\delta[\subseteq V\bigcap\mathbb{R}\subseteq V$. So if $x\in A\vdash 0<x-x_0<\delta$ then $x\in A\bigcap V\wedge x_0<x$ giving $|f(x)-L|<\varepsilon$.

$\Leftarrow$**.** Given $U\in\mathcal{T}_{\bar{\mathbb{R}}}\vdash L\in U$ then as $L\in\mathbb{R}$ we have $L\in U\bigcap\mathbb{R}\in\mathcal{T}_{\mathbb{R}}$ [using 17.36] so that $\exists\varepsilon>0$ such that $L\in]L-\varepsilon,L+\varepsilon[\subseteq U\bigcap\mathbb{R}$, by the hypothese there $\exists\delta>0$ such that if $x\in A\vdash 0<x-x_0<\delta$ we have $|f(x)-L|<\varepsilon\Rightarrow f(x)\in]L-\varepsilon,L+\varepsilon[\subseteq U\bigcap\mathbb{R}\subseteq U$. So if we take $V=]x_0-\delta,x_0+\delta[\in\mathcal{T}_{\mathbb{R}}\subseteq\mathcal{T}_{\bar{\mathbb{R}}}$ then $x_0\in V$ and if $x\in A\bigcap V$ with $x_0<x$ we have $0<x-x_0<\delta\wedge x\in A$, so that $f(x)\in U$. Proving that $\lim\limits_{x\uparrow x_0} f(x)=L$.

$\lim\limits_{x\rightarrow x_0} f(x)=L$**.**

$\Rightarrow$**.** Given $\varepsilon>0$ take $L\in]L-\varepsilon,L+\varepsilon[\in\mathcal{T}_{\mathbb{R}}\subseteq\mathcal{T}_{\bar{\mathbb{R}}}$ [as $L\in\mathbb{R}$] then there exists a $V\in\mathcal{T}_{\bar{\mathbb{R}}}$ with $x_0\in V$ such that $\forall x\in V\bigcap A$ with $x\neq x_0$ we have $f(x)\in]L-\varepsilon,L+\varepsilon[\Rightarrow|f(x)-L|<\varepsilon$. As $x_0\in\mathbb{R}$ we have $x_0\in V\bigcap\mathbb{R}\in\mathcal{T}_{\mathbb{R}}$ [using 17.36] there exists a $\delta>0$ such that $x_0\in]x_0-\delta,x_0+\delta[\subseteq V\bigcap\mathbb{R}\subseteq V$. So if $x\in A\vdash 0<|x-x_0|<\delta$ we have $x\in A\wedge x\in V$ which gives that $|f(x)-L|<\varepsilon$.

$\Leftarrow$. Given $U\in\mathcal{T}_{\bar{\mathbb{R}}}$ with $L\in U$ then as $L\in\mathbb{R}$ we have $L\in U\bigcap\mathbb{R}\in\mathcal{T}_{\mathbb{R}}$ [using 17.36] so there $\exists\varepsilon>0$ such that $L\in]L-\varepsilon,L+\varepsilon[\subseteq U\bigcap\mathbb{R}\subseteq U$, by the hypothese there exists a $\delta>0$ such that $\forall x\in A\vdash 0<|x_0-x|<\delta$ we have $|f(x)-L|<\varepsilon\Rightarrow f(x)\in]L-\varepsilon, L+\varepsilon[\subseteq U$. So if we take $V=]x_0-\delta,x_0+\delta[$ then $x_0\in V$ and if $x\in A\bigcap V$ with $x_0\neq x$ we have $0<|x_0-x|<\delta$, so that $f(x)\in U$. Proving that $\lim_{x\rightarrow x_0} f(x)=L$

$x_0\in\mathbb{R}, L=\infty$. then we have for the different limits the following alternative definitions

$\lim_{x\uparrow x_0} f(x)=\infty$.

$\Rightarrow$. Given $\varepsilon>0$ take $L\in]\varepsilon,\infty]\in\mathcal{T}_{\bar{\mathbb{R}}}$ then there exists a $V\in\mathcal{T}_{\bar{\mathbb{R}}}$ with $x_0\in V$ such that $\forall x\in A\bigcap V$ with $x<x_0$ we have $f(x)\in]\varepsilon,\infty]$. As $x_0\in\mathbb{R}$ we have $x_0\in V\bigcap\mathbb{R}\in\mathcal{T}_{\mathbb{R}}$ [using 17.36] so there exists a $\delta>0$ so that $x_0\in]x_0-\delta,x_0+\delta[\subseteq V\bigcap\mathbb{R}\subseteq V$. If then $x\in A\wedge 0<x_0-x<\delta$ we have $x\in A\bigcap V\wedge x<x_0$ so that $f(x)\in]\varepsilon,\infty]\Rightarrow\varepsilon<f(x)$.

$\Leftarrow$. Given $L\in U\in\mathcal{T}_{\bar{\mathbb{R}}}$ then as $L=\infty$ we have, using 17.36 and the fact that $\forall B\in\mathcal{B}_{\mathbb{R}}\bigcup\mathcal{B}_{-\infty}$ we have $\infty\notin B$, that $\exists\varepsilon>0$ such that $L\in]\varepsilon,\infty]\subseteq U$. Using the hypothese there exists a $\delta>0$ such that if $x\in A\vdash 0<x_0-x<\delta$ we have $\varepsilon<f(x)\Rightarrow f(x)\in U$. If we take then $V=]x_0-\delta,x_0+\delta[\in\mathcal{T}_{\bar{\mathbb{R}}}$ we have $x_0\in V$ and $\forall x\in V\bigcap A\vdash x<x_0$ we have $x\in A\wedge 0<x_0-x<\delta$ so that $f(x)\in U$. Proving that $\lim_{x\uparrow x_0} f(x)=\infty=L$.

$\lim_{x\downarrow x_0} f(x)=\infty$.

$\Rightarrow$. Given $\varepsilon>0$ take $L\in]\varepsilon,\infty]\in\mathcal{T}_{\bar{\mathbb{R}}}$ then there exists a $V\in\mathcal{T}_{\bar{\mathbb{R}}}$ with $x_0\in V$ such that $\forall x\in A\bigcap V$ with $x_0<x$ we have $f(x)\in]\varepsilon,\infty]$. As $x_0\in\mathbb{R}$ we have $x_0\in V\bigcap\mathbb{R}\in\mathcal{T}_{\mathbb{R}}$ [using 17.36] so there exists a $\delta>0$ so that $x_0\in]x_0-\delta,x_0+\delta[\subseteq V\bigcap\mathbb{R}\subseteq V$. If then $x\in A\wedge 0<x-x_0<\delta$ we have $x\in A\bigcap V\wedge x_0<x$ so that $f(x)\in]\varepsilon,\infty]\Rightarrow\varepsilon<f(x)$.

$\Leftarrow$. Given $L\in U\in\mathcal{T}_{\bar{\mathbb{R}}}$ then as $L=\infty$ we have, using 17.36 and the fact that $\forall B\in\mathcal{B}_{\mathbb{R}}\bigcup\mathcal{B}_{-\mathbb{R}}$ we have $\infty\notin B$, that $\exists\varepsilon>0$ such that $L\in]\varepsilon,\infty]\subseteq U$. Using the hypothese there exists a $\delta>0$ such that if $x\in A\vdash 0<x-x_0<\delta$ we have $\varepsilon<f(x)\Rightarrow f(x)\in U$. If we take then $V=]x_0-\delta,x_0+\delta[\in\mathcal{T}_{\bar{\mathbb{R}}}$ we have $x_0\in V$ and $\forall x\in V\bigcap A\vdash x_0<x$ we have $x\in A\wedge 0<x-x_0<\delta$ so that $f(x)\in U$. Proving that $\lim_{x\uparrow x_0} f(x)=\infty=L$.

$\lim_{x\rightarrow x_0} f(x)=\infty$.

$\Rightarrow$. Given $\varepsilon>0$ take $L\in]\varepsilon,\infty]\in\mathcal{T}_{\bar{\mathbb{R}}}$ then there exists a $V\in\mathcal{T}_{\bar{\mathbb{R}}}$ with $x_0\in V$ such that $\forall x\in A\bigcap V$ with $x\neq x_0$ we have $f(x)\in]\varepsilon,\infty]$. As $x_0\in\mathbb{R}$ we have $x_0\in V\bigcap\mathbb{R}\in\mathcal{T}_{\mathbb{R}}$ [using 17.36] so there exists a $\delta > 0$ so that $x_0 \in]x_0 - \delta, x_0 + \delta[\subseteq V \bigcap \mathbb{R} \subseteq V$. If then $x \in A \wedge 0 < |x_0 - x| < \delta$ we have $x \in A \bigcap V \wedge x \neq x_0$ so that $f(x) \in]\varepsilon, \infty] \Rightarrow \varepsilon < f(x)$.

$\Leftarrow$. Given $L \in U \in \mathcal{T}_{\bar{\mathbb{R}}}$ then as $L = \infty$ we have, using 17.36 and the fact that $\forall B \in \mathcal{B}_{\mathbb{R}} \bigcup \mathcal{B}_{-\infty}$ we have $\infty \notin B$, that $\exists \varepsilon > 0$ such that $L \in]\varepsilon, \infty] \subseteq U$. Using the hypothese there exists a $\delta > 0$ such that if $x \in A \vdash 0 < |x_0 - x| < \delta$ we have $\varepsilon < f(x) \Rightarrow f(x) \in U$. If we take then $V =]x_0 - \delta, x_0 + \delta[\in \mathcal{T}_{\bar{\mathbb{R}}}$ we have $x_0 \in V$ and $\forall x \in V \bigcap A \vdash x \neq x_0$ we have $x \in A \wedge 0 < |x_0 - x| < \delta$ so that $f(x) \in U$. Proving that $\lim_{x \uparrow x_0} f(x) = \infty = L$.

$x_0 \in \mathbb{R}, L = -\infty$. then we have for the different limits the following alternative definitions

$\lim_{x \uparrow x_0} f(x) = -\infty$.

$\Rightarrow$. Given $\varepsilon > 0$ take $L \in [-\infty, -\varepsilon[\in \mathcal{T}_{\bar{\mathbb{R}}}$ then there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $x_0 \in V$ such that $\forall x \in A \bigcap V$ with $x < x_0$ we have $f(x) \in [-\infty, -\varepsilon[$. As $x_0 \in \mathbb{R}$ we have $x_0 \in V \bigcap \mathbb{R} \in \mathcal{T}_{\mathbb{R}}$ [using 17.36] so there exists a $\delta > 0$ so that $x_0 \in]x_0 - \delta, x_0 + \delta[\subseteq V \bigcap \mathbb{R} \subseteq V$. If then $x \in A \wedge 0 < x_0 - x < \delta$ we have $x \in A \bigcap V \wedge x < x_0$ so that $f(x) \in [-\infty, \varepsilon[\Rightarrow f(x) < -\varepsilon$.

$\Leftarrow$. Given $L \in U \in \mathcal{T}_{\bar{\mathbb{R}}}$ then as $L = -\infty$ we have, using 17.36 and the fact that $\forall B \in \mathcal{B}_{\mathbb{R}} \bigcup \mathcal{B}_{\infty}$ we have $-\infty \notin B$, that $\exists \varepsilon > 0$ such that $L \in [-\infty, -\varepsilon[\subseteq U$. Using the hypothese there exists a $\delta > 0$ such that if $x \in A \vdash 0 < x_0 - x < \delta$ we have $f(x) < -\varepsilon \Rightarrow f(x) \in U$. If we take then $V =]x_0 - \delta, x_0 + \delta[\in \mathcal{T}_{\bar{\mathbb{R}}}$ we have $x_0 \in V$ and $\forall x \in V \bigcap A \vdash x < x_0$ we have $x \in A \wedge 0 < x_0 - x < \delta$ so that $f(x) \in U$. Proving that $\lim_{x \uparrow x_0} f(x) = -\infty = L$.

$\lim_{x \downarrow x_0} f(x) = -\infty$.

$\Rightarrow$. Given $\varepsilon > 0$ take $L \in [-\infty, -\varepsilon[\in \mathcal{T}_{\bar{\mathbb{R}}}$ then there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $x_0 \in V$ such that $\forall x \in A \bigcap V$ with $x_0 < x$ we have $f(x) \in [-\infty, -\varepsilon[$. As $x_0 \in \mathbb{R}$ we have $x_0 \in V \bigcap \mathbb{R} \in \mathcal{T}_{\mathbb{R}}$ [using 17.36] so there exists a $\delta > 0$ so that $x_0 \in]x_0 - \delta, x_0 + \delta[\subseteq V \bigcap \mathbb{R} \subseteq V$. If then $x \in A \wedge 0 < x - x_0 < \delta$ we have $x \in A \bigcap V \wedge x_0 < x$ so that $f(x) \in [-\infty, \varepsilon[\Rightarrow f(x) < -\varepsilon$.

$\Leftarrow$. Given $L \in U \in \mathcal{T}_{\bar{\mathbb{R}}}$ then as $L = -\infty$ we have, using 17.36 and the fact that $\forall B \in \mathcal{B}_{\mathbb{R}} \bigcup \mathcal{B}_{\infty}$ we have $-\infty \notin B$, that $\exists \varepsilon > 0$ such that $L \in [-\infty, -\varepsilon[\subseteq U$. Using the hypothese there exists a $\delta > 0$ such that if $x \in A \vdash 0 < x - x_0 < \delta$ we have $f(x) < -\varepsilon \Rightarrow f(x) \in U$. If we take then $V =]x_0 - \delta, x_0 + \delta[\in \mathcal{T}_{\bar{\mathbb{R}}}$ we have $x_0 \in V$ and $\forall x \in V \bigcap A \vdash x < x_0$ we have $x \in A \wedge 0 < x - x_0 < \delta$ so that $f(x) \in U$. Proving that $\lim_{x \uparrow x_0} f(x) = -\infty = L$.

$\boldsymbol{\lim_{x \rightarrow x_0} f(x) = -\infty}$**.**

⇒. Given $\varepsilon > 0$ take $L \in [-\infty, -\varepsilon[\in \mathcal{T}_{\bar{\mathbb{R}}}$ then there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $x_0 \in V$ such that $\forall x \in A \bigcap V$ with $x \neq x_0$ we have $f(x) \in [-\infty, -\varepsilon[$. As $x_0 \in \mathbb{R}$ we have $x_0 \in V \bigcap \mathbb{R} \in \mathcal{T}_{\mathbb{R}}$ [using 17.36] so there exists a $\delta > 0$ so that $x_0 \in]x_0 - \delta, x_0 + \delta[\subseteq V \bigcap \mathbb{R} \subseteq V$. If then $x \in A \wedge 0 < |x_0 - x| < \delta$ we have $x \in A \bigcap V \wedge x \neq x_0$ so that $f(x) \in [-\infty, \varepsilon[\Rightarrow f(x) < -\varepsilon$.

⇐. Given $L \in U \in \mathcal{T}_{\bar{\mathbb{R}}}$ then as $L = -\infty$ we have, using 17.36 and the fact that $\forall B \in \mathcal{B}_{\mathbb{R}} \bigcup \mathcal{B}_{\infty}$ we have $-\infty \notin B$, that $\exists \varepsilon > 0$ such that $L \in [-\infty, -\varepsilon[\subseteq U$. Using the hypothese there exists a $\delta > 0$ such that if $x \in A \vdash 0 < |x_0 - x| < \delta$ we have $f(x) < -\varepsilon \Rightarrow f(x) \in U$. If we take then $V =]x_0 - \delta, x_0 + \delta[\in \mathcal{T}_{\bar{\mathbb{R}}}$ we have $x_0 \in V$ and $\forall x \in V \bigcap A \vdash x \neq x_0$ we have $x \in A \wedge 0 < |x_0 - x| < \delta$ so that $f(x) \in U$. Proving that $\lim_{x \uparrow x_0} f(x) = -\infty = L$.

$\boldsymbol{x_0 = \infty, L \in \mathbb{R}}$**.** then we have for the different limits the following alternative definitions

$\boldsymbol{\lim_{x \uparrow \infty} f(x) = L}$**.**

⇒. If $\varepsilon > 0$ take then $L \in]L - \varepsilon, L + \varepsilon[\in \mathcal{T}_{\mathbb{R}} \subseteq \mathcal{T}_{\bar{\mathbb{R}}}$ [as $L \in \mathbb{R}$]. We have then that there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ such that if $x \in A \bigcap V \vdash x < \infty$ we have $f(x) \in]L - \varepsilon, L + \varepsilon[\Rightarrow |f(x) - L| < \varepsilon$. As $\forall B \in \mathcal{B}_{\mathbb{R}} \bigcup \mathcal{B}_{-\infty}$ we have $\infty \notin B$ we have that $\exists \delta > 0$ such that $\infty \in]\delta, \infty] \subseteq V$, so if $x \in A$ and $\delta < x < \infty$ we have $x \in A \bigcap V \vdash x < \infty$ and thus $|f(x) - L| < \varepsilon$

⇐. Let $L \in U \in \mathcal{T}_{\bar{\mathbb{R}}}$ then as $L \in \mathbb{R}$ we have $L \in U \bigcap \mathbb{R} \in \mathcal{T}_{\mathbb{R}}$ [using 17.36] so there exists a $\varepsilon > 0$ such that $L \in]L - \varepsilon, L + \varepsilon[\subseteq U \bigcap \mathbb{R} \subseteq U$. By the hypothese there exists a $\delta > 0$ such that if $\delta < x < \infty$ we have $|f(x) - L| < \varepsilon \Rightarrow f(x) \in]L - \varepsilon, L + \varepsilon[\subseteq U$. Take now $V =]\delta, \infty]$ then if $x \in V \bigcap A \vdash x < \infty$ we have $f(x) \in U$ proving that $\lim_{x \uparrow x_0} f(x) = L$

$\boldsymbol{\lim_{x \downarrow \infty} f(x) = L}$**.** then given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ we have for $\infty \in]1, \infty] \in \mathcal{T}_{\bar{\mathbb{R}}}$ that $\forall x \in]1, \infty] \vdash x > \infty$ that $f(x) \in U$ is satisfied vacuously, so this is always true.

$\boldsymbol{\lim_{x \rightarrow \infty} f(x) = L}$**.**

⇒. Given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ such that $\forall x \in V \bigcap A \vdash x \neq \infty$ we have $f(x) \in U$, so $\forall x \in V \bigcap A \vdash x < \infty$ we have $x \neq \infty$ giving that $f(x) \in U$, proving that $\lim_{x \uparrow \infty} f(x) = L$

⇐. Given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ such that $\forall x \in V \bigcap A \vdash x < \infty$ we have $f(x) \in U$, so $\forall x \in V \bigcap A \vdash x \neq \infty$ we have $x < \infty$ giving that $f(x) \in U$, proving that $\lim_{x \rightarrow \infty} f(x) = L$

$x_0 = \infty, L = \infty$. then we have for the different limits the following alternative definitions

$\lim_{x\uparrow\infty} f(x) = \infty$.

$\Rightarrow$. Given $\varepsilon > 0$ we have that $L = \infty \in]\varepsilon, \infty]$ so there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $\infty \in V$ such that $\forall x \in A \bigcap V \vdash x < \infty$ we have $f(x) \in]\varepsilon, \infty] \Rightarrow \varepsilon < f(x)$. As $\infty \in V$ there exists a $\delta > 0$ such that $\infty \in]\delta, \infty] \subseteq V$. So if $x \in A \vdash \delta < x < \infty$ then $\varepsilon < f(x)$

$\Leftarrow$. Given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ there exists a $\varepsilon > 0$ such that $\infty \in]\varepsilon, \infty] \subseteq U$, by the hypothese there exists then a $\delta > 0$ such that if $x \in A \wedge \delta < x < \infty$ then $\varepsilon < f(x) \Rightarrow f(x) \in U$. Take now $V =]\delta, \infty] \in \mathcal{T}_{\bar{\mathbb{R}}}$ then we have $\forall x \in A \bigcap V \vdash x < \infty$ that $f(x) \in U$

$\lim_{x\downarrow\infty} f(x) = \infty$. then given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ we have for $\infty \in]1, \infty] \in \mathcal{T}_{\bar{\mathbb{R}}}$ that $\forall x \in]1, \infty] \vdash x > \infty$ that $f(x) \in U$ is satisfied vacuously. So this is always true.

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

$\Rightarrow$. Given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ such that $\forall x \in V \bigcap A \vdash x \neq \infty$ we have $f(x) \in U$, so $\forall x \in V \bigcap A \vdash x < \infty$ we have $x \neq \infty$ giving that $f(x) \in U$, proving that $\lim_{x\uparrow\infty} f(x) = L$

$\Leftarrow$. Given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ such that $\forall x \in V \bigcap A \vdash x < \infty$ we have $f(x) \in U$, so $\forall x \in V \bigcap A \vdash x \neq \infty$ we have $x < \infty$ giving that $f(x) \in U$, proving that $\lim_{x\to\infty} f(x) = L$

$x_0 = \infty, L = -\infty$. then we have for the different limits the following alternative definitions

$\lim_{x\uparrow\infty} f(x) = -\infty$.

$\Rightarrow$. Given $\varepsilon > 0$ we have that $L = -\infty \in [-\infty, -\varepsilon[$ so there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $\infty \in V$ such that $\forall x \in A \bigcap V \vdash x < \infty$ we have $f(x) \in [-\infty, -\varepsilon[\Rightarrow f(x) < -\varepsilon$. As $\infty \in V$ there exists a $\delta > 0$ such that $\infty \in]\delta, \infty] \subseteq V$. So if $x \in A \vdash \delta < x < \infty$ then $f(x) < -\infty$

$\Leftarrow$. Given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ there exists a $\varepsilon > 0$ such that $-\infty \in [-\infty, -\varepsilon[\subseteq U$, by the hypothese there exists then a $\delta > 0$ such that if $x \in A \wedge \delta < x < \infty$ then $f(x) < -\varepsilon \Rightarrow f(x) \in U$. Take now $V =]\delta, \infty] \in \mathcal{T}_{\bar{\mathbb{R}}}$ then we have $\forall x \in A \bigcap V \vdash x < \infty$ that $f(x) \in U$

$\lim_{x\downarrow\infty} f(x) = -\infty$. then given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ we have for $\infty \in]1, \infty] \in \mathcal{T}_{\bar{\mathbb{R}}}$ that $\forall x \in]1, \infty] \vdash x > \infty$ that $f(x) \in U$ is satisfied vacuously. So this is always true.

$\lim_{x\to\infty} f(x) = -\infty$.

$\Rightarrow$. Given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ such that $\forall x \in V \bigcap A \vdash x < \infty$ we have $f(x) \in U$, so $\forall x \in V \bigcap A \vdash x \neq \infty$ we have $x < \infty$ giving that $f(x) \in U$, proving that $\lim_{x\uparrow\infty} f(x) = L$

$\Leftarrow$. Given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ such that $\forall x \in V \bigcap A \vdash x \neq \infty$ we have $f(x) \in U$, so $\forall x \in V \bigcap A \vdash x < \infty$ we have $x \neq \infty$ giving that $f(x) \in U$, proving that $\lim_{x \to \infty} f(x) = L$

$x_0 = -\infty, L \in \mathbb{R}$. then we have for the different limits the following alternative definitions

$\lim_{x \uparrow -\infty} f(x) = L$. then given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ we have for $-\infty \in [-\infty, -1[\in \mathcal{T}_{\bar{\mathbb{R}}}$ that $\forall x \in]1, \infty] \vdash x > \infty \;\; f(x) \in U$ is satisfied vacuously. So this is always true.

$\lim_{x \downarrow -\infty} f(x) = L$.

$\Rightarrow$. If $\varepsilon > 0$ take then $L \in]L - \varepsilon, L + \varepsilon[\in \mathcal{T}_{\mathbb{R}} \subseteq \mathcal{T}_{\bar{\mathbb{R}}}$ [as $L \in \mathbb{R}$]. We have then that there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ such that if $x \in A \bigcap V \vdash -\infty < x$ we have $f(x) \in]L - \varepsilon, L + \varepsilon[\Rightarrow |f(x) - L| < \varepsilon$. As $\forall B \in \mathcal{B}_{\mathbb{R}} \bigcup \mathcal{B}_{\infty}$ we have $-\infty \notin B$ we have that $\exists \delta > 0$ such that $-\infty \in [-\infty, -\delta[\subseteq V$, so if $x \in A$ and $-\infty < x < -\delta$ we have $x \in A \bigcap V \vdash -\infty < x$ and thus $|f(x) - L| < \varepsilon$

$\Leftarrow$. Let $L \in U \in \mathcal{T}_{\bar{\mathbb{R}}}$ then as $L \in \mathbb{R}$ we have $L \in U \bigcap \mathbb{R} \in \mathcal{T}_{\mathbb{R}}$ [using 17.36] so there exists a $\varepsilon > 0$ such that $L \in]L - \varepsilon, L + \varepsilon[\subseteq U \bigcap \mathbb{R} \subseteq U$. By the hypothese there exists a $\delta > 0$ such that if $-\infty < x < -\delta$ we have $|f(x) - L| < \varepsilon \Rightarrow f(x) \in]L - \varepsilon, L + \varepsilon[\subseteq U$. Take now $V = [-\infty, -\delta[$ then if $x \in V \bigcap A \vdash -\infty < x$ we have $f(x) \in U$ proving that $\lim_{x \uparrow x_0} f(x) = L$

$\lim_{x \to -\infty} f(x) = L$.

$\Rightarrow$. Given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ such that $\forall x \in V \bigcap A \vdash -\infty < x$ we have $f(x) \in U$, so $\forall x \in V \bigcap A \vdash x \neq -\infty$ we have $-\infty < x$ giving that $f(x) \in U$, proving that $\lim_{x \downarrow -\infty} f(x) = L$

$\Leftarrow$. Given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ such that $\forall x \in V \bigcap A \vdash -\infty < x$ we have $f(x) \in U$, so $\forall x \in V \bigcap A \vdash x \neq -\infty$ we have $-\infty < x$ giving that $f(x) \in U$, proving that $\lim_{x \to -\infty} f(x) = L$

$x_0 = -\infty, L = \infty$. then we have for the different limits the following alternative definitions

$\lim_{x \uparrow -\infty} f(x) = \infty$. then given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ we have for $-\infty \in [-\infty, -1[\in \mathcal{T}_{\bar{\mathbb{R}}}$ that $\forall x \in]1, \infty] \vdash x > \infty \;\; f(x) \in U$ is satisfied vacuously. So this is always true.

$\lim_{x \downarrow -\infty} f(x) = \infty$.

$\Rightarrow$. Given $\varepsilon > 0$ we have that $L = \infty \in]\varepsilon, \infty]$ so there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $-\infty \in V$ such that $\forall x \in A \bigcap V \vdash -\infty < x$ we have $f(x) \in]\varepsilon, \infty] \Rightarrow \varepsilon < f(x)$. As $-\infty \in V$ there exists a $\delta > 0$ such that $-\infty \in [-\infty, -\delta[\subseteq V$. So if $x \in A \vdash -\infty < x < -\delta$ then $\varepsilon < f(x)$

$\Leftarrow$. Given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ there exists a $\varepsilon > 0$ such that $\infty \in]\varepsilon, \infty] \subseteq U$, by the hypothese there exists then a $\delta > 0$ such that if $x \in A \wedge -\infty < x < -\delta$ then $\varepsilon < f(x) \Rightarrow f(x) \in U$. Take now $V = [-\infty, -\delta[\in \mathcal{T}_{\bar{\mathbb{R}}}$ then we have $\forall x \in A \bigcap V \vdash -\infty < x$ that $f(x) \in U$

$\lim_{x \to -\infty} f(x) = \infty$.

$\Rightarrow$. Given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ such that $\forall x \in V \bigcap A \vdash -\infty < x$ we have $f(x) \in U$, so $\forall x \in V \bigcap A \vdash x \neq -\infty$ we have $-\infty < x$ giving that $f(x) \in U$, proving that $\lim_{x \downarrow -\infty} f(x) = L$

$\Leftarrow$. Given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ such that $\forall x \in V \bigcap A \vdash -\infty < x$ we have $f(x) \in U$, so $\forall x \in V \bigcap A \vdash x \neq -\infty$ we have $-\infty < x$ giving that $f(x) \in U$, proving that $\lim_{x \to -\infty} f(x) = L$

$x_0 = -\infty, L = -\infty$. then we have for the different limits the following alternative definitions

$\lim_{x \uparrow -\infty} f(x) = -\infty$. then given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ we have for $-\infty \in [-\infty, -1[\in \mathcal{T}_{\bar{\mathbb{R}}}$ that $\forall x \in]1, \infty] \vdash x > \infty \quad f(x) \in U$ is satisfied vacuously. So this is always true.

$\lim_{x \downarrow -\infty} f(x) = -\infty$.

$\Rightarrow$. Given $\varepsilon > 0$ we have that $L = -\infty \in [-\infty. -\varepsilon[$ so there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $-\infty \in V$ such that $\forall x \in A \bigcap V \vdash -\infty < x$ we have $f(x) \in [-\infty, -\varepsilon[\Rightarrow f(x) < -\varepsilon$. As $-\infty \in V$ there exists a $\delta > 0$ such that $-\infty \in [-\infty, -\delta[\subseteq V$. So if $x \in A \vdash -\infty < x < -\delta$ then $f(x) < -\varepsilon$

$\Leftarrow$. Given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L = -\infty \in U$ there exists a $\varepsilon > 0$ such that $-\infty \in [-\infty, -\varepsilon[\subseteq U$, by the hypothese there exists then a $\delta > 0$ such that if $x \in A \wedge -\infty < x < -\delta$ then $f(x) < -\varepsilon \Rightarrow f(x) \in U$. Take now $V = [-\infty, -\delta[\in \mathcal{T}_{\bar{\mathbb{R}}}$ then we have $\forall x \in A \bigcap V \vdash -\infty < x$ that $f(x) \in U$

$\lim_{x \to -\infty} f(x) = -\infty$.

$\Rightarrow$. Given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ such that $\forall x \in V \bigcap A \vdash -\infty < x$ we have $f(x) \in U$, so $\forall x \in V \bigcap A \vdash x \neq -\infty$ we have $-\infty < x$ giving that $f(x) \in U$, proving that $\lim_{x \downarrow -\infty} f(x) = L$

$\Leftarrow$. Given $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ with $L \in U$ there exists a $V \in \mathcal{T}_{\bar{\mathbb{R}}}$ such that $\forall x \in V \bigcap A \vdash -\infty < x$ we have $f(x) \in U$, so $\forall x \in V \bigcap A \vdash x \neq -\infty$ we have $-\infty < x$ giving that $f(x) \in U$, proving that $\lim_{x \to -\infty} f(x) = L$ $\square$

Lemma 17.58. *Let $x \in \bar{\mathbb{R}}$ with $x \in U$ a open set in $\bar{\mathbb{R}}$ then we have*

1. *If $-\infty < x$ then there exists a $z \in \mathbb{R}$ such that $z < x$ and $]z, x] \subseteq U$*

2. *If* $x<\infty$ *then there exists a* $z\in\mathbb{R}$ *such that* $x<z$ *and* $[x,z[\subseteq U$

Note that if $x=\infty$ *or* $x=-\infty$ *we must have the inclusion* $]...x]$ *or* $[x...[$

Proof.

1. If $-\infty<x$ and U a open set in $\bar{\mathbb{R}}$ containing x then there exists a $B\in\mathcal{B}_{\bar{\mathbb{R}}}$ such that $x\in B\subseteq U$ now for B we have the following possibilities:

   $B=]y-\varepsilon,y+\varepsilon[$ where $y\in\mathbb{R}\wedge\varepsilon>0$. take then $z=y-\varepsilon$ then as $x\in B=]y-\varepsilon,y+\varepsilon[$ we have $z=y-\varepsilon<x$ and $]z,x]\subseteq]y-\varepsilon,y+\varepsilon[\subseteq U$.

   $B=[-\infty,y[$ where $y\in\mathbb{R}$. then as $x\in B$ we have $x<y$, from $-\infty<x<y<\infty$ we have then that $x\in\mathbb{R}$, take then $z=x-1$ then we have $z<x$ and $]z,x]\subseteq[-\infty,y[\subseteq U$.

   $B=]y,\infty]$ where $y\in\mathbb{R}$. then as $x\in B$ we have $y<x$ take then $z=y$ then we have $z<x$ and $]z,x]\subseteq]y,\infty]\subseteq U$

2. If $x<\infty$ and U a open set in $\bar{\mathbb{R}}$ containing x then there exists a $B\in\mathcal{B}_{\bar{\mathbb{R}}}$ such that $x\in B\subseteq U$ now for B we have the following possibilities:

   $B=]y-\varepsilon,y+\varepsilon[$ where $y\in\mathbb{R}\wedge\varepsilon>0$. take then $z=y+\varepsilon$ then as $x\in B=]y-\varepsilon,y+\varepsilon[$ we have $x<y+\varepsilon=z$ and $[x,z[\subseteq]y-\varepsilon,y+\varepsilon[\subseteq U$.

   $B=[-\infty,y[$ where $y\in\mathbb{R}$. then as $x\in B$ we have $x<y$, take then $z=y$ then we have $x<z$ and $[x,z[\subseteq[-\infty,y[\subseteq U$.

   $B=]y,\infty]$ where $y\in\mathbb{R}$. then as $x\in B$ we have $y<x$, from $-\infty<y<x<\infty$ we have $x\in\mathbb{R}$, take then $z=x+1$ then we have $x<z$ and $[x,z[\subseteq]y,\infty]\subseteq U$ □

Theorem 17.59. *Let* $A,B\subseteq\bar{\mathbb{R}}$, $f\colon A\rightarrow B$ *a function and* $x_0\in\bar{\mathbb{R}}$ *then we have*

1. *If f is a increasing function and x_0 is a left limit point of A then* $\lim_{x\uparrow x_0}f(x)=\sup(\{f(x)|x\in A\wedge x<x_0\})$
2. *If f is a increasing function and x_0 is a right limit point of A then* $\lim_{x\downarrow x_0}f(x)=\inf(\{f(x)|x\in A\wedge x_0<x\})$
3. *If f is a decreasing function and x_0 is a left limit point of A then* $\lim_{x\uparrow x_0}f(x)=\inf(\{f(x)|x\in A\wedge x<x_0\})$
4. *If f is a decreasing function and x_0 is a right limit point of A then* $\lim_{x\downarrow x_0}f(x)=\sup(\{f(x)|x\in A\wedge x_0<x\})$

Proof.

1. As x_0 is a left limit point we must have by 17.44 that $-\infty<x_0$. Take $L=\sup(\{f(x)|x\in A\wedge x<x_0\})$ then we have the following cases to consider

   $\forall x\in A\vdash x<x_0\vDash f(x)=-\infty$. then using 17.49 we have $\lim_{x\uparrow x_0}f(x)=-\infty=\sup(\{-\infty\})=\sup(\{f(x)|x\in A\wedge x<x_0\})=L$ so that $\lim_{x\uparrow x_0}f(x)=L$

$\exists x_1 \in A \vdash x_1 < x_0 \vDash f(x_1) > -\infty$. then $L > -\infty$. If now U is a open set in $\bar{\mathbb{R}}$ with $L \in U$, then using the previous lemma (see 17.58) there exists a $z \in \mathbb{R}$ such that $z < L$ and $]z, L] \subseteq U$. From the definition of a supremum and $z < L$ it follows that there exists a $y \in A$ with $y < x_0$ such that $z < f(y) \leqslant L$. We have now to consider the following two cases for x_0

$x_0 = \infty$. take then the open set $V =]y, x_0 + 1]$ containing x_0, if now $x \in V \bigcap A$ with $x < x_0$ then $y < x$ so as f is increasing we have $z < f(y) \leqslant f(x) \leqslant L \Rightarrow f(x) \in]z, L] \subseteq U$

$x_0 \in \mathbb{R}$. for y we have as $y < x_0$ two possibilities

$y = -\infty$. take the open set $V =]x_0 - 1, y]$ containing x_0, if now $x \in V \bigcap A$ with $x_0 < x$ then $x \leqslant y$ so as f is increasing we have $L \leqslant f(x) \leqslant f(y) < z \Rightarrow f(x) \in [L, z[\subseteq U$

$y \in \mathbb{R}$. define then $\delta = x_0 - y$ and take the open set $V =]x_0 - \delta, x_0 + \delta[$, if now $x \in V \bigcap A$ with $x_0 < x$ then $y = x_0 - \delta < x$ so as f is increasing we have $z < f(y) \leqslant f(x) \leqslant L \Rightarrow f(x) \in]z, L] \subseteq U$

This proves that given U open with $L \in U$ there exists a V open containing x_0 such that $\forall x \in V \bigcap A$ with $x < x_0$ we have $f(x) \in U$ proving that $\lim_{x \uparrow x_0} f(x) = L$

2. As x_0 is a right limit point we must have by 17.44 that $x_0 < \infty$. Take $L = \inf(\{f(x) | x \in A \wedge x_0 < x\})$ then we have the following cases to consider

$\forall x \in A \vdash x_0 < x$ we have $f(x) = \infty$. then using 17.49 we have $\lim_{x \downarrow x_0} f(x) = \inf(\{\infty\}) = \inf(\{f(x) | x \in A \wedge x_0 < x\}) = L$ so that $\lim_{x \downarrow x_0} f(x) = L$

$\exists x_1 \in A \vdash x_0 < x_1 \vDash f(x) < \infty$. then $L < \infty$. If now U is a open set in $\bar{\mathbb{R}}$ with $L \in U$, then using the previous lemma (see 17.58) there exists a $z \in \bar{\mathbb{R}}$ such that $L < z$ and $[L, z[\subseteq U$. From the definition of a infinum and $L < z$ it follows that there exists a $y \in A$ with $x_0 < y$ such that $L \leqslant f(y) < z$. We have now to consider the following two cases for x_0

$x_0 = -\infty$. take then the open set $V = [-\infty, y[$ containing x_0, if now $x \in V \bigcap A$ with $x_0 < x$ then $x < y$ so as f is increasing we have $L \leqslant f(x) \leqslant f(y) < z \Rightarrow f(x) \in [L, z[\subseteq U$

$x_0 \in \mathbb{R}$. for y we have as $x_0 < y$ two possibilities

$y = \infty$. take the open set $V =]x_0 - 1, y]$ containing x_0, if now $x \in V \bigcap A$ with $x_0 < x$ then $x \leqslant y$ so as f is increasing we have $L \leqslant f(x) \leqslant f(y) < z \Rightarrow f(x) \in [L, z[\subseteq U$

$y \in \mathbb{R}$. define then $\delta = y - x_0$ and take the open set $V =]x_0 - \delta, x_0 + \delta[$, if now $x \in V \bigcap A$ with $x_0 < x$ then $x < x_0 + \delta = y$ so as f is increasing we have $L \leqslant f(x) \leqslant f(y) < z \Rightarrow f(x) \in [L, z[\subseteq U$

This proves that given U open with $L \in U$ there exists a V open containing x_0 such that $\forall x \in V \bigcap A$ with $x_0 < x$ we have $f(x) \in U$ proving that $\lim_{x \downarrow x_0} f(x) = L$

3. As x_0 is a left limit point we must have by 17.44 that $-\infty < x_0$. Take $L = \inf(\{f(x)|x \in A \wedge x < x_0\})$ then we have the following cases to consider

   $\forall x \in A \vdash x < x_0 \vDash f(x) = \infty$. then using 17.49 we have $\lim_{x \uparrow x_0} f(x) = \infty = \inf(\{\infty\}) = \inf(\{f(x)|x \in A \wedge x < x_0\}) = L$ so that $\lim_{x \uparrow x_0} f(x) = L$

   $\exists x_1 \in A \vdash x_1 < x_0 \vDash f(x_1) < \infty$. then $L < \infty$. If now U is a open set in $\bar{\mathbb{R}}$ with $L \in U$, then using the previous lemma (see 17.58) there exists a $z \in \mathbb{R}$ such that $L < z$ and $[L, z[\subseteq U$. From the definition of a infinum and $L < z$ it follows that there exists a $y \in A$ with $y < x_0$ such that $L \leqslant f(y) < z$. We have now to consider the following two cases for x_0

   $x_0 = \infty$. take then the open set $V =]y, x_0]$ containing x_0, if now $x \in V \bigcap A$ with $x < x_0$ then $y < x$ so as f is decreasing we have $L \leqslant f(x) \leqslant f(y) < z \Rightarrow f(x) \in [L, z[\subseteq U$

   $x_0 \in \mathbb{R}$. for y we have as $y < x_0$ two possibilities

   $y = -\infty$. take the open set $V = [y, x_0 + 1[$ containing x_0, if now $x \in V \bigcap A$ with $x < x_0$ then $y \leqslant x$ so as f is decreasing we have $L \leqslant f(x) \leqslant f(y) < z \Rightarrow f(x) \in [L, z[\subseteq U$

   $y \in \mathbb{R}$. define then $\delta = x_0 - y$ and take the open set $V =]x_0 - \delta, x_0 + \delta[$, if now $x \in V \bigcap A$ with $x_0 < x$ then $y = x_0 - \delta < x$ so as f is decreasing we have $L \leqslant f(x) \leqslant f(y) < z \Rightarrow f(x) \in [L, x[\subseteq U$

   This proves that given U open with $L \in U$ there exists a V open containing x_0 such that $\forall x \in V \bigcap A$ with $x < x_0$ we have $f(x) \in U$ proving that $\lim_{x \uparrow x_0} f(x) = L$.

4. As x_0 is a right limit point we must have by 17.44 that $x_0 < \infty$. Take $L = \sup(\{f(x)|x \in A \wedge x_0 < x\})$ then we have the following cases to consider

   $\forall x \in A \vdash x < x_0 \vDash f(x) = -\infty$. then using 17.49 we have $\lim_{x \downarrow x_0} f(x) = -\infty = \sup(\{\infty\}) = \sup(\{f(x)|x \in A \wedge x_0 < x\}) = L$ so that $\lim_{x \downarrow x_0} f(x) = L$

   $\exists x_1 \in A \vdash x_0 < x_1 \vDash -\infty < f(x_1)$. then $-\infty < L$. If now U is a open set in $\bar{\mathbb{R}}$ with $L \in U$, then using the previous lemma (see 17.58) there exists a $z \in \mathbb{R}$ such that $z < L$ and $]z, L] \subseteq U$. From the definition of a supremum and $z < L$ it follows that there exists a $y \in A$ with $x_0 < y$ such that $z < f(y) \leqslant L$. We have now to consider the following two cases for x_0

   $x_0 = -\infty$. take then the open set $V = [x_0, y[$ containing x_0, if now $x \in V \bigcap A$ with $x_0 < x$ then $x < y$ so as f is decreasing we have $z < f(y) \leqslant f(x) \leqslant L \Rightarrow f(x) \in]z, L] \subseteq U$

$\boldsymbol{x_0 \in \mathbb{R}}$**.** for y we have as $x_0 < y$ two possibilities

$\boldsymbol{y = \infty}$**.** take the open set $V =]x_0 - 1, y]$ containing x_0, if now $x \in V \bigcap A$ with $x_0 < x$ then $x \leqslant y$ so as f is decreasing we have $z < f(y) \leqslant f(x) \leqslant L \Rightarrow f(x) \in]z, L] \subseteq U$

$\boldsymbol{y \in \mathbb{R}}$**.** define then $\delta = y - x_0$ and take the open set $V =]x_0 - \delta, x_0 + \delta[$, if now $x \in V \bigcap A$ with $x < x_0$ then $x < x_0 + \delta = y$ so as f is decreasing we have $z < f(y) \leqslant f(x) \leqslant L \Rightarrow f(x) \in]z, L] \subseteq U$

This proves that given U open with $L \in U$ there exists a V open containing x_0 such that $\forall x \in V \bigcap A$ with $x < x_0$ we have $f(x) \in U$ proving that $\lim_{x \downarrow x_0} f(x) = L$. □

Applying the above on limits at ∞ and $-\infty$ gives the following corollary.

Corollary 17.60. *Let $B \subseteq \bar{\mathbb{R}}$ then we have that*

1. *If $f: [-\infty, a] \rightarrow B$ is a increasing function where $a \in \mathbb{R} \bigcup \{\infty\}$ then* $\lim_{x \rightarrow -\infty} f(x) = \lim_{x \downarrow -\infty} f(x) = \inf(\{f(x) | -\infty < x \leqslant a\})$
2. *If $f: [-\infty, a[\rightarrow B$ is a increasing function where $a \in \mathbb{R} \bigcup \{\infty\}$ then* $\lim_{x \rightarrow -\infty} f(x) = \lim_{x \downarrow -\infty} f(x) = \inf(\{f(x) | -\infty < x < a\})$
3. *If $f:]-\infty, a] \rightarrow B$ is a increasing function where $a \in \mathbb{R} \bigcup \{\infty\}$ then* $\lim_{x \rightarrow -\infty} f(x) = \lim_{x \downarrow -\infty} f(x) = \inf(\{f(x) | -\infty < x \leqslant a\})$
4. *If $f:]-\infty, a[\rightarrow B$ is a increasing function where $a \in \mathbb{R} \bigcup \{\infty\}$ then* $\lim_{x \rightarrow -\infty} f(x) = \lim_{x \downarrow -\infty} f(x) = \inf(\{f(x) | -\infty < x < a\})$
5. *If $f: [-\infty, a] \rightarrow B$ is a decreasing function where $a \in \mathbb{R} \bigcup \{\infty\}$ then* $\lim_{x \rightarrow -\infty} f(x) = \lim_{x \downarrow -\infty} f(x) = \sup(\{f(x) | -\infty < x \leqslant a\})$
6. *If $f: [-\infty, a[\rightarrow B$ is a decreasing function where $a \in \mathbb{R} \bigcup \{\infty\}$ then* $\lim_{x \rightarrow -\infty} f(x) = \lim_{x \downarrow -\infty} f(x) = \sup(\{f(x) | -\infty < x < a\})$
7. *If $f:]-\infty, a] \rightarrow B$ is a decreasing function where $a \in \mathbb{R} \bigcup \{\infty\}$ then* $\lim_{x \rightarrow -\infty} f(x) = \lim_{x \downarrow -\infty} f(x) = \sup(\{f(x) | -\infty < x \leqslant a\})$
8. *If $f:]-\infty, a[\rightarrow B$ is a decreasing function where $a \in \mathbb{R} \bigcup \{\infty\}$ then* $\lim_{x \rightarrow -\infty} f(x) = \lim_{x \downarrow -\infty} f(x) = \sup(\{f(x) | -\infty < x < a\})$
9. *If $f: [a, \infty] \rightarrow B$ is a increasing function where $a \in \mathbb{R} \bigcup \{-\infty\}$ then* $\lim_{x \rightarrow \infty} f(x) = \lim_{x \uparrow \infty} f(x) = \sup(\{f(x) | a \leqslant x < \infty\})$
10. *If $f: [a, \infty[\rightarrow B$ is a increasing function where $a \in \mathbb{R} \bigcup \{-\infty\}$ then* $\lim_{x \rightarrow \infty} f(x) = \lim_{x \uparrow \infty} f(x) = \sup(\{f(x) | a \leqslant x < \infty\})$

11. *If* $f:]a, \infty] \rightarrow B$ *is a increasing function where* $a \in \mathbb{R} \bigcup \{-\infty\}$ *then* $\lim_{x \rightarrow \infty} f(x) = \lim_{x \uparrow \infty} f(x) = \sup(\{f(x) | a < x < \infty\})$

12. *If* $f:]a, \infty[\rightarrow B$ *is a increasing function where* $a \in \mathbb{R} \bigcup \{-\infty\}$ *then* $\lim_{x \rightarrow \infty} f(x) = \lim_{x \uparrow \infty} f(x) = \sup(\{f(x) | a < x < \infty\})$

13. *If* $f: [a, \infty] \rightarrow B$ *is a decreasing function where* $a \in \mathbb{R} \bigcup \{-\infty\}$ *then* $\lim_{x \rightarrow \infty} f(x) = \lim_{x \uparrow \infty} f(x) = \inf(\{f(x) | a \leqslant x < \infty\})$

14. *If* $f: [a, \infty[\rightarrow B$ *is a decreasing function where* $a \in \mathbb{R} \bigcup \{-\infty\}$ *then* $\lim_{x \rightarrow \infty} f(x) = \lim_{x \uparrow \infty} f(x) = \inf(\{f(x) | a \leqslant x < \infty\})$

15. *If* $f:]a, \infty] \rightarrow B$ *is a decreasing function where* $a \in \mathbb{R} \bigcup \{-\infty\}$ *then* $\lim_{x \rightarrow \infty} f(x) = \lim_{x \uparrow \infty} f(x) = \inf(\{f(x) | a < x < \infty\})$

16. *If* $f:]a, \infty[\rightarrow B$ *is a decreasing function where* $a \in \mathbb{R} \bigcup \{-\infty\}$ *then* $\lim_{x \rightarrow \infty} f(x) = \lim_{x \uparrow \infty} f(x) = \inf(\{f(x) | a < x < \infty\})$

Proof. This follows from the fact that in cases 1,2,3,4,5,7,8 $-\infty$ is a left limit point of the domain of f, in cases $9, 10, 11, 12, 13, 14, 15, 16$ is ∞ is a right limit point of the domain of f, theorem 17.54 and the previous theorem 17.59. $\square$

Theorem 17.61. *Let* $A, B \subseteq \bar{\mathbb{R}}$, $x_0 \in \bar{\mathbb{R}}$, $x_0 \in W$ *a open set and* $f: A \rightarrow B$, $g: A \rightarrow B$ *be functions so that* $\forall x \in A \bigcap W$ *we have* $f(x) \leqslant g(x)$ *then we have*

1. *If* $\lim_{x \downarrow x_0} f(x)$, $\lim_{x \downarrow x_0} g(x)$ *exists then we have* $\lim_{x \downarrow x_0} f(x) \leqslant \lim_{x \downarrow x_0} g(x)$
2. *If* $\lim_{x \uparrow x_0} f(x)$, $\lim_{x \uparrow x_0} g(x)$ *exists then we have* $\lim_{x \uparrow x_0} f(x) \leqslant \lim_{x \uparrow x_0} g(x)$
3. *If* $\lim_{x \rightarrow x_0} f(x)$, $\lim_{x \rightarrow x_0} g(x)$ *exists then we have* $\lim_{x \rightarrow x_0} f(x) \leqslant \lim_{x \rightarrow x_0} g(x)$

Proof. The proofs for the the different cases are similar

1. Take $L_f = \lim_{x \downarrow x_0} f(x)$, $L_g = \lim_{x \downarrow x_0} g(x)$ and assume that $L_g < L_f$, the following possibilities exists then for L_f, L_g

   $\boldsymbol{L_g \in \mathbb{R} \wedge L_f = \infty}$**.** Then $L_g \in]L_g - 1, L_g + 1[$ a open set so by the definition of the right limit there exists a open set V_g containing x_0 so that $\forall x \in V_g \bigcap A$ with $x_0 < x$ we have $L_g - 1 < g(x) < L_g + 1$. Also $]L_g + 1, \infty]$ is a open set containing $L_f = \infty$ so by the definition of a right limit there exists a open set V_f containing x_0 so that $\forall x \in V_f \bigcap A$ with $x_0 < x$ we have $L_g + 1 < f(x) \leqslant \infty$. Now as x_0 is a right limit point and $x_0 \in (V_f \bigcap V_g) \bigcap W$ there exists a $y_0 \in ((V_f \bigcap V_g) \bigcap W) \bigcap A$ with $x_0 < y_0$ and thus by the above we have $g(y_0) < L_g + 1 < f(y_0)$ contradicting the fact that $f(y_0) \leqslant g(y_0)$. So we must have $L_f \leqslant L_g$

$L_g = -\infty \wedge L_f = \infty$. Then $-\infty = L_g \in [-\infty, 1[$ a open set so by definition of the right limit there exists a open set V_g containing x_0 so that $\forall x \in V_g \bigcap A$ with $x_0 < x$ we have $-\infty \leqslant g(x) < 1$. Also $]1, \infty]$ is a open set containing $L_f = \infty$ so by the definition of a right limit there exists a open set V_f containing x_0 so that $\forall x \in V_f \bigcap A$ with $x_0 < x$ we have $1 < f(x) \leqslant \infty$. Now as x_0 is a right limit point and $y_0 \in ((V_f \bigcap V_g) \bigcap W) \bigcap A$ there exists a $y_0 \in (V_f \bigcap V_g) \bigcap A$ with $x_0 < y_0$ and thus by the above we have $g(y_0) < 1 < f(y_0)$ contradicting the fact that $f(y_0) \leqslant g(y_0)$. So we must have $L_f \leqslant L_g$

$L_g \in \mathbb{R} \wedge L_f \in \mathbb{R}$. Take then $\varepsilon = \frac{L_f - L_g}{2}$. We have then that $L_g \in]L_g - \varepsilon, L_g + \varepsilon[$ a open set so by definition of the right limit there exists a open set V_g containing x_0 so that $\forall x \in V_g \bigcap A$ with $x_0 < x$ we have $L_g - \varepsilon < g(x) < L_g + \varepsilon = \frac{L_f + L_g}{2}$. Also $]L_f - \varepsilon, L_f + \varepsilon[$ is a open set containing L_f thus we have by the definition of a right limit that there exists a open set V_f such that $\forall x \in V_f \bigcap A$ with $x_0 < x$ we have $\frac{L_f + L_g}{2} = L_f - \varepsilon < f(x) < L_f + \varepsilon$. Now as x_0 is a right limit point and $x_0 \in (V_f \bigcap V_g) \bigcap W$ there exists a $y_0 \in ((V_f \bigcap V_g) \bigcap W) \bigcap A$ with $x_0 < y_0$ and thus by the above we have $g(y_0) < \frac{L_f + L_g}{2} < f(y_0)$ contradicting the fact that $f(y_0) \leqslant g(y_0)$. So we must have $L_f \leqslant L_g$.

$L_g = -\infty \wedge L_f \in \mathbb{R}$. Then $-\infty = L_g \in [-\infty, L_f - 1[$ a open set so by the definition of the right limit there exists a open set V_g containing x_0 so that $\forall x \in V_g \bigcap A$ with $x_0 < x$ we have $-\infty \leqslant g(x) < L_f - 1$. Also $]L_f - 1, L_f + 1[$ is a open set containing L_f so by the definition of a right limit there exists a open set V_f such that $\forall x \in V_f \bigcap A$ with $x_0 < x$ we have $L_f - 1 < f(x) < L_f + 1$. Now as x_0 is a right limit point and $x_0 \in (V_f \bigcap V_g) \bigcap W$ there exists a $y_0 \in ((V_f \bigcap V_g) \bigcap W) \bigcap A$ with $x_0 < y_0$ and thus by the above we have $g(y_0) < L_f - 1 < f(y_0)$ contradicting the fact that $f(y_0) \leqslant g(y_0)$. So we must have $L_f \leqslant L_g$.

2. Take $L_f = \lim_{x \uparrow x_0} f(x)$, $L_g = \lim_{x \downarrow x_0} g(x)$ and assume that $L_g < L_f$, the following possibilities exists then for L_f, L_g

$L_g \in \mathbb{R} \wedge L_f = \infty$. Then $L_g \in]L_g - 1, L_g + 1[$ a open set so by the definition of the left limit there exists a open set V_g containing x_0 so that $\forall x \in V_g \bigcap A$ with $x < x_0$ we have $L_g - 1 < g(x) < L_g + 1$. Also $]L_g + 1, \infty]$ is a open set containing $L_f = \infty$ so by the definition of a left limit there exists a open set V_f containing x_0 so that $\forall x \in V_f \bigcap A$ with $x < x_0$ we have $L_g + 1 < f(x) \leqslant \infty$. Now as x_0 is a left limit point and $x_0 \in (V_f \bigcap V_g) \bigcap W$ there exists a $y_0 \in ((V_f \bigcap V_g) \bigcap W) \bigcap A$ with $y_0 < x_0$ and thus by the above we have $g(y_0) < L_g + 1 < f(y_0)$ contradicting the fact that $f(y_0) \leqslant g(y_0)$. So we must have $L_f \leqslant L_g$

$L_g = -\infty \wedge L_f = \infty$. Then $-\infty = L_g \in [-\infty, 1[$ a open set so by definition of the left limit there exists a open set V_g containing x_0 so that $\forall x \in V_g \bigcap A$ with $x < x_0$ we have $-\infty \leqslant g(x) < 1$. Also $]1, \infty]$ is a open set containing $L_f = \infty$ so by the definition of a left limit there exists a open set V_f containing x_0 so that $\forall x \in V_f \bigcap A$ with

$x < x_0$ we have $1 < f(x) \leqslant \infty$. Now as x_0 is a left limit point and $x_0 \in (V_f \bigcap V_g) \bigcap W$ there exists a $y_0 \in ((V_f \bigcap V_g) \bigcap W) \bigcap A$ with $y_0 < x_0$ and thus by the above we have $g(y_0) < 1 < f(y_0)$ contradicting the fact that $f(y_0) \leqslant g(y_0)$. So we must have $L_f \leqslant L_g$

$\boldsymbol{L_g \in \mathbb{R} \wedge L_f \in \mathbb{R}}$**.** Take then $\varepsilon = \frac{L_f - L_g}{2}$. We have then that $L_g \in]L_g - \varepsilon, L_g + \varepsilon[$ a open set so by definition of the left limit there exists a open set V_g containing x_0 so that $\forall x \in V_g \bigcap A$ with $x < x_0$ we have $L_g - \varepsilon < g(x) < L_g + \varepsilon = \frac{L_f + L_g}{2}$. Also $]L_f - \varepsilon, L_f + \varepsilon[$ is a open set containing L_f thus we have by the definition of a left limit that there exists a open set V_f such that $\forall x \in V_f \bigcap A$ with $x < x_0$ we have $\frac{L_f + L_g}{2} = L_f - \varepsilon < f(x) < L_f + \varepsilon$. Now as x_0 is a left limit point and $x_0 \in (V_f \bigcap V_g) \bigcap W$ there exists a $y_0 \in ((V_f \bigcap V_g) \bigcap W) \bigcap A$ with $y_0 < x_0$ and thus by the above we have $g(y_0) < \frac{L_f + L_g}{2} < f(y_0)$ contradicting the fact that $f(y_0) \leqslant g(y_0)$. So we must have $L_f \leqslant L_g$.

$\boldsymbol{L_g = -\infty \wedge L_f \in \mathbb{R}}$**.** Then $-\infty = L_g \in [-\infty, L_f - 1[$ a open set so by the definition of the left limit there exists a open set V_g containing x_0 so that $\forall x \in V_g \bigcap A$ with $x < x_0$ we have $-\infty \leqslant g(x) < L_f - 1$. Also $]L_f - 1, L_f + 1[$ is a open set containing L_f so by the definition of a left limit there exists a open set V_f such that $\forall x \in V_f \bigcap A$ with $x < x_0$ we have $L_f - 1 < f(x) < L_f + 1$. Now as x_0 is a left limit point and $x_0 \in (V_f \bigcap V_g) \bigcap W$ there exists a $y_0 \in ((V_f \bigcap V_g) \bigcap W) \bigcap A$ with $y_0 < x_0$ and thus by the above we have $g(y_0) < L_f - 1 < f(y_0)$ contradicting the fact that $f(y_0) \leqslant g(y_0)$. So we must have $L_f \leqslant L_g$.

3. Take $L_f = \lim_{x \to x_0} f(x)$, $L_g = \lim_{x \to x_0} g(x)$ and assume that $L_g < L_f$, the following possibilities exists then for L_f, L_g

$\boldsymbol{L_g \in \mathbb{R} \wedge L_f = \infty}$**.** Then $L_g \in]L_g - 1, L_g + 1[$ a open set so by the definition of the limit there exists a open set V_g containing x_0 so that $\forall x \in V_g \bigcap A$ with $x_0 \neq x$ we have $L_g - 1 < g(x) < L_g + 1$. Also $]L_g + 1, \infty]$ is a open set containing $L_f = \infty$ so by the definition of a limit there exists a open set V_f containing x_0 so that $x_0 \in (V_f \bigcap V_g) \bigcap W$ with $x_0 \neq x$ we have $L_g + 1 < f(x) \leqslant \infty$. Now as x_0 is a limit point and $x_0 \in V_f \bigcap V_g$ there exists a $y_0 \in ((V_f \bigcap V_g) \bigcap W) \bigcap A$ with $x_0 \neq y_0$ and thus by the above we have $g(y_0) < L_g + 1 < f(y_0)$ contradicting the fact that $f(y_0) \leqslant g(y_0)$. So we must have $L_f \leqslant L_g$

$\boldsymbol{L_g = -\infty \wedge L_f = \infty}$**.** Then $-\infty = L_g \in [-\infty, 1[$ a open set so by definition of the limit there exists a open set V_g containing x_0 so that $\forall x \in V_g \bigcap A$ with $x_0 \neq x$ we have $-\infty \leqslant g(x) < 1$. Also $]1, \infty]$ is a open set containing $L_f = \infty$ so by the definition of a limit there exists a open set V_f containing x_0 so that $\forall x \in V_f \bigcap A$ with $x_0 \neq x$ we have $1 < f(x) \leqslant \infty$. Now as x_0 is a limit point and $x_0 \in (V_f \bigcap V_g) \bigcap W$ there exists a $y_0 \in ((V_f \bigcap V_g) \bigcap W) \bigcap A$ with $x_0 \neq y_0$ and thus by the above we have $g(y_0) < 1 < f(y_0)$ contradicting the fact that $f(y_0) \leqslant g(y_0)$. So we must have $L_f \leqslant L_g$

$L_g \in \mathbb{R} \wedge L_f \in \mathbb{R}$. Take then $\varepsilon = \frac{L_f - L_g}{2}$. We have then that $L_g \in]L_g - \varepsilon, L_g + \varepsilon[$ a open set so by definition of the limit there exists a open set V_g containing x_0 so that $\forall x \in V_g \bigcap A$ with $x_0 \neq x$ we have $L_g - \varepsilon < g(x) < L_g + \varepsilon = \frac{L_f + L_g}{2}$. Also $]L_f - \varepsilon, L_f + \varepsilon[$ is a open set containing L_f thus we have by the definition of a limit that there exists a open set V_f such that $\forall x \in V_f \bigcap A$ with $x_0 \neq x$ we have $\frac{L_f + L_g}{2} = L_f - \varepsilon < f(x) < L_f + \varepsilon$. Now as x_0 is a limit point and $x_0 \in (V_f \bigcap V_g) \bigcap W$ there exists a $y_0 \in ((V_f \bigcap V_g) \bigcap W) \bigcap A$ with $x_0 \neq y_0$ and thus by the above we have $g(y_0) < \frac{L_f + L_g}{2} < f(y_0)$ contradicting the fact that $f(y_0) \leqslant g(y_0)$. So we must have $L_f \leqslant L_g$.

$L_g = -\infty \wedge L_f \in \mathbb{R}$. Then $-\infty = L_g \in [-\infty, L_f - 1[$ a open set so by the definition of the limit there exists a open set V_g containing x_0 so that $\forall x \in V_g \bigcap A$ with $x_0 \neq x$ we have $-\infty \leqslant g(x) < L_f - 1$. Also $]L_f - 1, L_f + 1[$ is a open set containing L_f so by the definition of a limit there exists a open set V_f such that $\forall x \in V_f \bigcap A$ with $x_0 \neq x$ we have $L_f - 1 < f(x) < L_f + 1$. Now as x_0 is a limit point and $x_0 \in (V_f \bigcap V_g) \bigcap W$ there exists a $y_0 \in ((V_f \bigcap V_g) \bigcap W) \bigcap A$ with $x_0 \neq y_0$ and thus by the above we have $g(y_0) < L_f - 1 < f(y_0)$ contradicting the fact that $f(y_0) \leqslant g(y_0)$. So we must have $L_f \leqslant L_g$. □

Corollary 17.62. *Let $A, B \subseteq \bar{\mathbb{R}}$, $x_0, L \in \bar{\mathbb{R}}$, $x_0 \in W$ a open set and $f: A \rightarrow B$ then*

1. *If $\forall x \in A \bigcap W$ we have $f(x) \leqslant L$ then*

   a. *If $\lim_{x \downarrow x_0} f(x)$ exists then $\lim_{x \downarrow x_0} f(x) \leqslant L$*

   b. *If $\lim_{x \uparrow x_0} f(x)$ exists then $\lim_{x \uparrow x_0} f(x) \leqslant L$*

   c. *If $\lim_{x \rightarrow x_0} f(x)$ exists then $\lim_{x \rightarrow x_0} f(x) \leqslant L$*

2. *If $\forall x \in A \bigcap W$ we have $L \leqslant f(x)$ then*

   a. *If $\lim_{x \downarrow x_0} f(x)$ exists then $L \leqslant \lim_{x \downarrow x_0} f(x)$*

   b. *If $\lim_{x \uparrow x_0} f(x)$ exists then $L \leqslant \lim_{x \uparrow x_0} f(x)$*

   c. *If $\lim_{x \rightarrow x_0} f(x)$ exists then $L \leqslant \lim_{x \rightarrow x_0} f(x)$*

Proof.

1. In this define then $g: A \rightarrow B$ by $x \rightarrow L$ so that $\forall x \in A \bigcap W$ we have $f(x) \leqslant L = g(x)$, applying then the previous theorem (see 17.61) gives

   a. $\lim_{x \downarrow x_0} f(x) \leqslant \lim_{x \downarrow x_0} g(x) \underset{17.49}{=} L$

   b. $\lim_{x \uparrow x_0} f(x) \leqslant \lim_{x \uparrow x_0} g(x) \underset{17.49}{=} L$

c. $\lim_{x \rightarrow x_0} f(x) \leqslant \lim_{x \rightarrow x_0} g(x) \underset{17.49}{=} L$

2. In this define then $g: A \rightarrow B$ by $x \rightarrow L$ so that $\forall x \in A \bigcap W$ we have $g(x) = L \leqslant f(x)$, applying then the previous theorem (see 17.61) gives

   a. $L \underset{17.49}{=} \lim_{x \downarrow x_0} g(x) \leqslant \lim_{x \downarrow x_0} f(x)$

   b. $L \underset{17.49}{=} \lim_{x \uparrow x_0} g(x) \leqslant \lim_{x \uparrow x_0} f(x)$

   c. $L \underset{17.49}{=} \lim_{x \rightarrow x_0} g(x) \leqslant \lim_{x \rightarrow x_0} f(x)$ □

17.2.2 Sequences in $\bar{\mathbb{R}}$

Note that the sum in $\bar{\mathbb{R}}$ is not always defined so if we want to see how the supremum and infinum behaves on sums we have to be carefull, that is the reason that we restrict ourself to positve extended reals.

Notation 17.63. *Given $k \in \mathbb{N}$ we note $\{i \in \mathbb{N} | k \leqslant i\}$ as $\{k, ..., \infty\}$*

Definition 17.64. *Let $\{x_i\}_{i \in \{k, ..., \infty\}}$ be a sequence in a partial ordered set $\langle S, \leqslant \rangle$ then we say that $\{x_i\}_{i \in \{k, ..., \infty\}}$ is a increasing (decreasing) sequence if $\forall i \in \{k, ..., \infty\}$ we have $x_i \leqslant x_{i+1}$ $(x_{i+1} \leqslant x_i)$*

Lemma 17.65. *Let $\{x_i\}_{i \in \{k, ..., \infty\}}$ be a increasing (or decreasing) sequence then $\forall n, m \in \{k, ..., \infty\}$ with $n \leqslant m$ we have $x_n \leqslant x_m$ $(x_m \leqslant x_n)$*

Proof. The proof is trivial by mathematical induction, let $\{x_i\}_{i \in \{k, ..., \infty\}}$ be a increasing (or decreasing) sequence $n \in \{k, ..., \infty\}$ and define $S_n = \{m \in \mathbb{N}_0 | x_n \leqslant x_{n+m} (\text{or } x_{n+m} \leqslant x_n)\}$ then we have

$0 \in S$. this is trivial as $x_n \leqslant x_n = x_{n+0}$ (or $x_{n+0} = x_n \leqslant x_n$)

if $m \in S \Rightarrow m + 1 \in S$. from $m \in S$ we have $x_n \leqslant x_{n+m} \leqslant x_{n+m+1} \Rightarrow x_n \leqslant x_{n+m+1}$ (or $x_{n+m+1} \leqslant x_{n+m} \leqslant x_n \Rightarrow x_{n+m+1} \leqslant x_n$) proving $m + 1 \in S$

Let now $n, m \in \{k, ..., \infty\}$ with $n \leqslant m$ (or $m \leqslant n$) then $i = m - n \in S_n \Rightarrow x_n \leqslant x_{n+(m-n)} = x_m$ (or $x_m = x_{n+(m-n)} \leqslant x_n$) □

Recap the following result for increasing/decreasing sequences of real numbers (see 12.354)

Lemma 17.66. *Let $\{x_i\}_{i \in \{k, ..., \infty\}}$ be a increasing (decreasing) sequence of real numbers then we have*

1. *If $\{x_i | i \in \{k, ..., \infty\}\}$ has a supremum (infinum) then $\lim_{i \rightarrow \infty} x_i$ exists and $\lim_{i \rightarrow \infty} x_i = \sup(\{x_i | i \in \{k, ..., \infty\}\})$*
2. *If $\lim_{i \rightarrow \infty} x_i$ exists then $\sup(\{x_i | i \in \{k, ..., \infty\}\})$ $(\inf(\{x_i | i \in \{k, ..., \infty\}\}))$ exists and $\sup(\{x_i | i \in \{k, ..., \infty\}\}) = \lim_{i \rightarrow \infty} x_i$ $(\inf(\{x_i | i \in \{k, ..., \infty\}\}) = \lim_{i \rightarrow \infty} x_i)$*

Proof. This is already proved in 12.354 □

Theorem 17.67. *Let $\{x_i\}_{i\in\{n,\ldots,\infty\}}$ be a sequence of extended reals then*

1. $\inf(\{\sup(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{n,\ldots,\infty\}\})$ *exists*
2. $\sup(\{\inf(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{n,\ldots,\infty\}\})$ *exists*
3. $\forall m\geqslant n$ *we have*

   $$\inf(\{\sup(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{n,\ldots,\infty\}\})=\inf(\{\sup(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{m,\ldots,\infty\}\})$$

4. $\forall m\geqslant n$ *we have*

   $$\sup(\{\inf(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{n,\ldots,\infty\}\})=\sup(\{\inf(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{m,\ldots,\infty\}\})$$

5. $\forall m\in\mathbb{N}$ *we have*

   a. $\inf(\{\sup(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{n,\ldots,\infty\}\})=\inf(\{\sup(\{x_{i+m}|i\in\{k,\ldots,\infty\}\})|k\in\{n,\ldots,\infty\}\})$

   b. $\sup(\{\inf(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{n,\ldots,\infty\}\})=\sup(\{\inf(\{x_{i+n}|i\in\{k,\ldots,\infty\}\})|k\in\{n,\ldots,\infty\}\})$

Proof.

1. The existence follows from the fact that in $\bar{\mathbb{R}}$ every non empty set has a sup and inf (see 17.10)

2. The existence follows from the fact that in $\bar{\mathbb{R}}$ every non empty set has a sup and inf (see 17.10)

3. Given $k\in\{n,\ldots,\infty\}$ take $S_k=\sup(\{x_i|i\in\{k,\ldots.,\infty\}\})$ then $\{S_k|k\in\{m,\ldots,\infty\}\}\subseteq\{S_k|k\in\{n,\ldots,\infty\}\}$ proving using 2.171 that $\inf(S_k|k\in\{n,\ldots,,\infty\})\leqslant\inf(S_k|k\in\{m,\ldots,\infty\})$ proving that

   $$\inf(\{\sup(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{n,\ldots,\infty\}\})\leqslant\inf(\{\sup(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{m,\ldots,\infty\}\}) \tag{17.8}$$

   For the opposite inclusion, if $k\in\{n,\ldots,\infty\}$ then take then $l=\max(k,m)$ then $k\leqslant l$ and $l\in\{m,\ldots,\infty\}$. We have then that $\{x_i|i\in\{l,\ldots,\infty\}\}\subseteq\{x_i|i\in\{k,\ldots,\infty\}\}$ so that using 2.171 $S_l=\sup(\{x_i|i\in\{l,\ldots,\infty\}\})\leqslant\sup(\{x_i|i\in\{k,\ldots,\infty\}\})=S_k$. So for every $S_k\in\{S_i|i\in\{n,\ldots,\infty\}\}$ we find a $S_l\in\{S_i|i\in\{m,\ldots,\infty\}\}$ such that $S_l\leqslant S_k$. Using 2.172 we have then that $\inf(S_i|i\in\{m,\ldots,\infty\})\leqslant\inf(S_i|i\in\{n,\ldots,\infty\})$ or $\inf(\{\sup(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{m,\ldots,\infty\}\})\leqslant\inf(\{\sup(\{x_i|i\in\{n,\ldots,\infty\}\})|k\in\{n,\ldots,\infty\}\})$. Hence using 17.8 we have

   $$\inf(\{\sup(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{n,\ldots,\infty\}\})=\inf(\{\sup(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{m,\ldots,\infty\}\})$$

4. Given $k\in\{n,\ldots,\infty\}$ take $I_i=\inf(\{x_i|i\in\{k,\ldots,\infty\}\})$ then $\{I_i|i\in\{m,\ldots,\infty\}\}\subseteq\{I_i|i\in\{n,\ldots,\infty\}\}$ proving using 2.171 that $\sup(\{S_k|k\in\{m,\ldots,\infty\}\})\leqslant\sup(\{S_k|k\in\{n,\ldots,\infty\}\})$ proving that

$$\sup(\{\inf(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{m,\ldots,\infty\}\}\leqslant\sup(\{\inf(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{n,\ldots,\infty\}\}) \tag{17.9}$$

For the opposite inclusion, ikf $k\in\{n,\ldots,\infty\}$ take then $l=\max(k,m)$ then $k\leqslant l$ and $l\in\{m,.,.,\infty\}$. We have then that $\{x_i|i\in\{l,\ldots,\infty\}\}\subseteq\{x_i|i\in\{k,\ldots,\infty\}\}$ so that using 2.171 we have $I_k=\inf(\{x_i|i\in\{k,\ldots,\infty\}\})\leqslant\inf(\{x_i|i\in\{l,\ldots,\infty\}\})=I_l$. So for every $I_k\in\{I_i|i\in\{n,\ldots,\infty\}\}$ we can find a $I_l\in\{I_i|i\in\{m,\ldots,\infty\}\}$ such that $I_k\leqslant I_l$. Using 2.172 we have then that $\sup(\{I_i|i\in\{n,\ldots,\infty\}\})\leqslant\sup(\{I_i|i\in\{m,\ldots,\infty\}\})$ or $\sup(\{\inf(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{n,\ldots,\infty\}\})\leqslant\sup(\{\inf(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{m,\ldots,\infty\}\})$. Hence using 17.9 we have that

$$\sup(\{\inf(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{n,\ldots,\infty\}\})=\sup(\{\inf(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{m,\ldots,\infty\}\})$$

5. Given $k\in\{n+m,\ldots,\infty\}$ define $A_k=\{x_i|i\in\{k,\ldots,\infty\}\}$ and for $k\in\{n,\ldots,\infty\}$ $B_k=\{x_{i+m}|i\in\{k,\ldots,\infty\}\}$. Then we have for $k\in\{n+m,\ldots,\infty\}$

$$\begin{aligned}
x\in A_k &\Rightarrow \exists i\in\{k,\ldots,\infty\}\text{ such that }x=x_i\\
&\Rightarrow k\leqslant i\Rightarrow k-m\leqslant i-m\wedge x_i=x_{(i-m)+m}\\
&\Rightarrow x_i\in B_{k-m}\\
&\Rightarrow x\in B_{k-m}\\
x\in B_{k-m} &\Rightarrow \exists i\in\{k-m,\ldots,\infty\}\text{ such that }x=x_{i+m}\\
&\Rightarrow k-m\leqslant i\Rightarrow k\leqslant i+m\\
&\Rightarrow x\in A_k
\end{aligned}$$

proving that

$$\forall k\in\{n+m,\ldots,\infty\}\text{ we have that }A_k=B_{k-m} \tag{17.10}$$

On the other hand if $k\in\{n,\ldots,\infty\}$ then we have

$$\begin{aligned}
x\in B_k &\Rightarrow \exists i\in\{k,\ldots,\infty\}\text{ such that }x=x_{i+m}\\
&\Rightarrow k+m\leqslant i+m\\
&\Rightarrow x\in A_{k+m}\\
x\in A_{k+m} &\Rightarrow \exists i\in\{k+m,\ldots,\infty\}\text{ such that }x=x_i\\
&\Rightarrow k+m\leqslant i\Rightarrow k\leqslant i-m\Rightarrow x_{(i-m)+m}\\
&\Rightarrow x\in B_k
\end{aligned}$$

proving that

$$\forall k\in\{n,\ldots,\infty\}\text{ we have that }B_k=A_{k+m} \tag{17.11}$$

Now if $x \in \{\sup(A_k) | k \in \{n+m, \ldots, \infty\}\}$ then $\exists k \in \{n+m, \ldots, \infty\}$ such that $x = \sup(A_k) \underset{17.10}{=} \sup(B_{k-m})$where $k - m \in \{n, \ldots \infty\}$, hence there exists a $y = \sup(B_{k-m}) \in \{\sup(B_k) | k \in \{n, \ldots., \infty\}\}$ such that $x = y$ which using 2.172 proves that

$$\inf(\{\sup(B_k) | k \in \{n, \ldots, \infty\}\}) \leqslant \inf(\{\sup(A_k) | k \in \{n+m, \ldots, \infty\}\}) \tag{17.12}$$

If $x \in \{\sup(B_k) | k \in \{n, \ldots, \infty\}\}$ then $\exists k \in \{1, \ldots, n\}$ such that $x = \sup(B_k) \underset{17.11}{=} \sup(A_{k+m})$ where $k+m \in \{n+m, \ldots, \infty\}$, hence there exists a $y = \sup(A_{k+m}) \in \{\sup(A_k) | k \in \{n+m, \ldots, \infty\}\}$ such that $x = y$ which using 2.172 proves that

$$\inf(\{\sup(A_k) | k \in \{n+m, \ldots, \infty\}\}) \leqslant \inf(\{\sup(B_k) | k \in \{n, \ldots, \infty\}\}) \tag{17.13}$$

Combining 17.12 and 17.13 together with the definition of A_k, B_k we have

$$\inf(\{\sup(\{x_i | i \in \{k, \ldots, \infty\}\}) | k \in \{n+m, \ldots, \infty\}\}) = \inf(\{\sup(x_{i+m} | i \in \{k, \ldots, \infty\}) | k \in \{n, \ldots, \infty\}\}) \tag{17.14}$$

Using then (3) on the above proves

$$\inf(\{\sup(\{x_i | i \in \{k, \ldots, \infty\}\}) | k \in \{n, \ldots, \infty\}\}) = \inf(\{\sup(x_{i+m} | i \in \{k, \ldots, \infty\}) | k \in \{n, \ldots, \infty\}\}) \tag{17.15}$$

Now if $x \in \{\inf(A_k) | k \in \{n+m, \ldots, \infty\}\}$ then $\exists k \in \{n+m, \ldots, \infty\}$ such that $x = \inf(A_k) \underset{17.10}{=} \inf(B_{k-m})$where $k - m \in \{n, \ldots \infty\}$, hence there exists a $y = \sup(B_{k-m}) \in \{\inf(B_k) | k \in \{n, \ldots., \infty\}\}$ such that $x = y$ which using 2.172 proves that

$$\sup(\{\inf(A_k) | k \in \{n+m, \ldots, \infty\}\}) \leqslant \sup(\{\inf(B_k) | k \in \{n, \ldots, \infty\}\}) \tag{17.16}$$

If $x \in \{\inf(B_k) | k \in \{n, \ldots, \infty\}\}$ then $\exists k \in \{1, \ldots, n\}$ such that $x = \inf(B_k) \underset{17.11}{=} \inf(A_{k+m})$ where $k+m \in \{n+m, \ldots, \infty\}$, hence there exists a $y = \inf(A_{k+m}) \in \{\inf(A_k) | k \in \{n+m, \ldots, \infty\}\}$ such that $x = y$ which using 2.172 proves that

$$\sup(\{\inf(B_k) | k \in \{n, \ldots, \infty\}\}) \leqslant \sup(\{\inf(A_k) | k \in \{n+m, \ldots, \infty\}\}) \tag{17.17}$$

Combining 17.16 and 17.17 together with the definition of A_k, B_k we have

$$\inf(\{\sup(\{x_i | i \in \{k, \ldots, \infty\}\}) | k \in \{n+m, \ldots, \infty\}\}) = \inf(\{\sup(x_{i+m} | i \in \{k, \ldots, \infty\}) | k \in \{n, \ldots, \infty\}\}) \tag{17.18}$$

Using then (3) on the above proves

$$\inf(\{\sup(\{x_i | i \in \{k, \ldots, \infty\}\}) | k \in \{n, \ldots, \infty\}\}) = \inf(\{\sup(x_{i+m} | i \in \{k, \ldots, \infty\}) | k \in \{n, \ldots, \infty\}\}) \tag{17.19}$$

Finally 5 (a) is proved by 17.15 and 5 (b) by 17.19 □

The above theorem motivates the following definition

Definition 17.68. *Let $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ be a sequence of extended reals then*

$$\liminf_{i\to\infty} x_i = \sup\left(\{\inf\left(\{x_i|i\in\{l,\ldots,\infty\}\}\right)|l\in\{k,\ldots,\infty\}\}\right)$$

and

$$\limsup_{i\to\infty} x_i = \inf\left(\sup\left(\{x_i|i\in\{l,\ldots,\infty\}\}|l\in\{k,\ldots,\infty\}\right)\right).$$

Note 17.69. In the above above definition the starting index k is not mentioned in the notation $\limsup_{i\to\infty} x_i$ and $\liminf_{i\to\infty} x_i$ but this is ok because the previous theorem (see 17.67) we can use very number in $\{k,\ldots,\infty\}$ as the starting index.

Using this definition and the previous theorem (see 17.67 we have the following:

Theorem 17.70. *Let $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ be a sequence of extended reals then we have*

1. $\liminf_{i\to\infty} x_i$ *exists*
2. $\limsup_{i\to\infty} x_i$ *exists*
3. $\forall n\in\mathbb{N}$ *we have that*
   a. $\liminf_{i\to\infty} x_i = \liminf_{i\to\infty} x_{i+n}$
   b. $\limsup_{i\to\infty} x_i = \limsup_{i\to\infty} x_{i+n}$

Theorem 17.71. *Let $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ be a sequence of extended reals, $x\in\mathbb{R}$, such that $\{x_i+x\}_{i\in\{k,\ldots,\infty\}}$ is well defined then*

1. $\liminf_{i\to\infty}(x_i+x) = \left(\liminf_{i\to\infty} x_i\right)+x$
2. $\limsup_{i\to\infty}(x_i+x) = \left(\limsup_{i\to\infty} x_i\right)+x$

Proof.

1.

$$\begin{aligned}\liminf_{i\to\infty}(x_i+x) &= \sup\left(\{\inf\left(\{x_i+x|i\in\{l,\ldots,\infty\}\}\right)|l\in\{k,\ldots,\infty\}\}\right)\\ &\underset{17.29}{=} \sup\left(\{\inf\left(\{x_i|i\in\{l,\ldots,\infty\}\}\right)+x|l\in\{k,\ldots,\infty\}\}\right)\\ &\underset{17.29}{=} \sup\left(\{\inf\left(\{x_i|i\in\{l,\ldots,\infty\}\}\right)|l\in\{k,\ldots,\infty\}\}\right)+x\\ &= \left(\liminf_{i\to\infty} x_i\right)+x\end{aligned}$$

2.

$$\begin{aligned}\limsup_{i\to\infty}(x_i+x) &= \inf\left(\{\sup\left(\{x_i+x|i\in\{l,\ldots,\infty\}\}\right)|l\in\{k,\ldots,\infty\}\}\right)\\ &\underset{17.29}{=} \inf\left(\{\sup\left(\{x_i|i\in\{l,\ldots,\infty\}\}\right)+x|l\in\{,\ldots,\infty\}\}\right)\\ &\underset{17.29}{=} \inf\left(\{\sup\left(\{x_i|i\in\{l,\ldots,\infty\}\}\right)|l\in\{k,\ldots,\infty\}\}\right)+x\\ &= \left(\limsup_{i\to\infty} x_i\right)+x\end{aligned}$$

□

Theorem 17.72. *Let $\{x_i\}_{i\in\{k,\ldots,n\}}$ be a sequence of extended reals then*

1. $\lim_{i\to\infty}\inf(-x_i) = -\lim_{i\to\infty}\sup x_i$
2. $\lim_{i\to\infty}\sup(-x_i) = -\lim_{i\to\infty}\inf x_i$

Proof.

1.

$$\begin{aligned}\lim_{i\to\infty}\inf(-x_i) &= \sup(\{\inf(\{-x_i|i\in\{l,\ldots,\infty\}\})|l\in\{k,\ldots,\infty\}\}) \\ &\underset{17.30}{=} \sup(\{-\sup(\{x_i|i\in\{l,\ldots,\infty\}\})|l\in\{k,\ldots,\infty\}\}) \\ &\underset{17.30}{=} -\inf(\{\sup(\{x_i|i\in\{l,\ldots,\infty\}\})|l\in\{k,\ldots,\infty\}\}) \\ &= -\lim_{i\to\infty}\sup x_i\end{aligned}$$

2.

$$\begin{aligned}\lim_{i\to\infty}\sup(-x_i) &\underset{(1)}{=} -\lim_{i\to\infty}\inf(-(-x_i)) \\ &= -\lim_{i\to\infty}\inf x_i\end{aligned}$$

□

Lemma 17.73. *Let $a\in\bar{\mathbb{R}}$, $\{x_i\}_{i\in\{k,\ldots,\infty\}}, \{y_i\}_{i\in\{k,\ldots,\infty\}}\subseteq\mathbb{R}$ sequences of* ***reals*** *then*

1. *If $\forall i\in\{k,\ldots,\infty\}$ we have $x_i\leqslant x_{i+1}$ and $y_i\leqslant y_{i+1}$ then* $\sup(\{x_i|i\in\{k,\ldots,\infty\}\})+\sup(\{y_i|i\in\{k,\ldots,\infty\}\})=\sup(\{x_i+y_i|i\in\{k,\ldots,\infty\}\})$
2. *f $\forall i\in\{k,\ldots,\infty\}$ we have $x_{i+1}\leqslant x_i$ and $y_{i+1}\leqslant y_i$ then* $\inf(\{x_i|i\in\{k,\ldots,\infty\}\})+\inf(\{y_i|i\in\{k,\ldots,\infty\}\})=\inf(\{x_i+y_i|i\in\{k,\ldots,\infty\}\})$

Proof.

1. As $\forall i\in\{k,\ldots,\infty\}$ we have $x_i\leqslant\sup(\{x_j|j\in\{k,\ldots,\infty\}\})$ and $y_i\leqslant\sup(\{y_j|j\in\{k,\ldots,\infty\}\})$ so that $\forall i\in\{k,\ldots,\infty\}$ that $x_i+y_i\leqslant\sup(\{x_j|j\in\{k,\ldots,\infty\}\})+\sup(\{y_j|j\in\{k,\ldots,\infty\}\})$ hence

   $$\sup(\{x_i+y_i|i\in\{k,\ldots,\infty\}\})\leqslant\sup(\{x_i|i\in\{k,\ldots,\infty\}\})+\sup(\{y_i|i\in\{k,\ldots,\infty\}\}) \tag{17.20}$$

   Take now $l\in\{k,\ldots,\infty\}$ then for $i\in\{k,\ldots,\infty\}$ we have either

   $i<l$. then $x_l+y_i\leqslant x_l+y_l\leqslant\sup(\{x_j+y_j|j\in\{k,\ldots,\infty\}\})$ so that $y_i\leqslant\sup(\{x_j+y_j|j\in\{k,\ldots,\infty\}\})-x_l$

$l \leqslant i$. $x_l + y_i \leqslant x_i + y_i \leqslant \sup(\{x_j + y_j | j \in \{k, \ldots, \infty\}\})$ so that $y_i \leqslant \sup(\{x_j + y_j | j \in \{k, \ldots, \infty\}\}) - x_l$

as in all cases we have that $y_i \leqslant y_i \leqslant \sup(\{x_j + y_j | j \in \{k, \ldots, \infty\}\}) - x_l$ it follows that

$$\forall l \in \{k, \ldots, \infty\} \text{ we have } \sup(\{y_i | i \in \{k, \ldots, \infty\}\}) \leqslant \sup(\{x_j + y_j | j \in \{k, \ldots, \infty\}\}) - x_l \tag{17.21}$$

Consider now the fo;llowing cases for $\sup(\{y_i | i \in \{k, \ldots, \infty\}\})$

$\sup(\{y_i | i \in \{k, \ldots, \infty\}\}) = \infty$. then by 17.21 we have $\sup(\{x_j + y_j | j \in \{k, \ldots, \infty\}\}) = \infty$ so that $\sup(\{x_i | i \in \{k, \ldots, \infty\}\}) + \sup(\{y_i | i \in \{k, \ldots, \infty\}\}) \leqslant \sup(\{x_j + y_j | j \in \{k, \ldots, \infty\}\})$.

$\sup(\{y_i | i \in \{k, \ldots, \infty\}\}) \in \mathbb{R}$. then by 17.21 we have $\forall l \in \{k, \ldots, \infty\}$ that $x_l \leqslant \sup(\{x_j + y_j | j \in \{k, \ldots, \infty\}\}) - \sup(\{y_i | i \in \{k, \ldots, \infty\}\})$ proving that $\sup(\{x_i | i \in \{k, \ldots, \infty\}\}) \leqslant \sup(\{x_j + y_j | j \in \{k, \ldots, \infty\}\}) - \sup(\{y_i | i \in \{k, \ldots, \infty\}\})$ and thus $\sup(\{x_i | i \in \{k, \ldots, \infty\}\}) + \sup(\{y_i | i \in \{k, \ldots, \infty\}\}) \leqslant \sup(\{x_j + y_j | j \in \{k, \ldots, \infty\}\})$

As in all cases we have $\sup(\{x_i | i \in \{k, \ldots, \infty\}\}) + \sup(\{y_i | i \in \{k, \ldots, \infty\}\}) \leqslant \sup(\{x_j + y_j | j \in \{k, \ldots, \infty\}\})$ it follows using 17.20 that

$$\sup(\{x_i | i \in \{k, \ldots, \infty\}\}) + \sup(\{y_i | i \in \{k, \ldots, \infty\}\}) = \sup(\{x_j + y_j | j \in \{k, \ldots, \infty\}\})$$

2. Define $\{\bar{x}_i\}_{i \in \{k, \ldots, n\}}, \{\bar{y}_i\}_{i \in \{k, \ldots, \infty\}}$ by $\bar{x}_i = -x_i$ and $\bar{y}_i = -y_i$ then $\forall i \in \{k, \ldots, \infty\}$ we have $\bar{x}_i \leqslant \bar{x}_{i+1}$ and $\bar{y}_i \leqslant \bar{y}_{i+1}$. Hence using (1) we have that $\sup(\{\bar{x}_i | i \in \{k, \ldots, \infty\}\}) + \sup(\{\bar{y}_i | i \in \{k, \ldots, \infty\}\}) = \sup(\{\bar{x}_i + \bar{y}_i | i \in \{k, \ldots, \infty\}\})$ or $\sup(\{-x_i | i \in \{k, \ldots, \infty\}\}) + \sup(\{-y_i | i \in \{k, \ldots, \infty\}\}) = \sup(\{-(x_i + y_i) | i \in \{k, \ldots, \infty\}\})$. Using 17.30 we have then that

$$\inf(\{x_i | i \in \{k, \ldots, \infty\}\}) + \inf(\{y_i | i \in \{k, \ldots, \infty\}\}) = \inf(\{x_i + y_i | i \in \{k, \ldots, \infty\}\})$$

□

Theorem 17.74. $\liminf\limits_{i \to \infty}$ *and* $\limsup\limits_{i \to \infty}$ *have the following properties*

1. *If* $\{x_i\}_{i \in \{k, \ldots, \infty\}} \subseteq \bar{\mathbb{R}}$ *is a sequence of extended reals and* $0 \leqslant \alpha \in \mathbb{R}$ *then*

   a. $\liminf\limits_{i \to \infty} (\alpha \cdot x_i) = \alpha \cdot \liminf\limits_{i \to \infty} x_i$

   b. $\limsup\limits_{i \to \infty} (\alpha \cdot x_i) = \alpha \cdot \limsup\limits_{i \to \infty} x_i$

2. *If* $\{x_i\}_{i \in \{k, \ldots, n\}}, \{y_i\}_{i \in \{k, \ldots, n\}} \subseteq \mathbb{R}$ *be sequences of reals such that* $\left(\liminf\limits_{i \to \infty} x_i\right) + \left(\liminf\limits_{i \to \infty} y_i\right)$ *is well defined*

   a. $\left(\liminf\limits_{i \to \infty} x_i\right) + \left(\liminf\limits_{i \to \infty} y_i\right) \leqslant \liminf\limits_{i \to \infty} (x_i + y_i)$

b. $\limsup_{i\to\infty}(x_i+y_i)\leqslant\left(\liminf_{i\to\infty}x_i\right)+\left(\liminf_{i\to\infty}y_i\right)$

Proof.

1.

$$\begin{aligned}
\liminf_{i\to\infty}(\alpha_i\cdot x) &= \sup\left(\{\inf\left(\{\alpha\cdot x_i|i\in\{l,\ldots,\infty\}\}\right)|l\in\{k,\ldots,\infty\}\}\right)\\
&\underset{17.30}{=} \sup\left(\{\alpha\cdot\inf\left(\{x_i|i\in\{l,\ldots,\infty\}\}\right)|l\in\{k,\ldots,\infty\}\}\right)\\
&\underset{17.30}{=} \alpha\cdot\sup\left(\{\inf\left(\{x_i|i\in\{l,\ldots,\infty\}\}\right)|l\in\{k,\ldots,\infty\}\}\right)\\
&= \alpha\cdot\liminf_{i\to\infty}x_i\\
\limsup_{i\to\infty}(\alpha_i\cdot x) &= \inf\left(\{\sup\left(\{\alpha\cdot x_i|i\in\{l,\ldots,\infty\}\}\right)|l\in\{k,\ldots,\infty\}\}\right)\\
&\underset{17.30}{=} \inf\left(\{\alpha\cdot\sup\left(\{x_i|i\in\{l,\ldots,\infty\}\}\right)|l\in\{k,\ldots,\infty\}\}\right)\\
&\underset{17.30}{=} \alpha\cdot\inf\left(\{\sup\left(\{x_i|i\in\{l,\ldots,\infty\}\}\right)|l\in\{k,\ldots,\infty\}\}\right)\\
&= \alpha\cdot\limsup_{i\to\infty}x_i
\end{aligned}$$

2.

a. Let $l\in\{k,\ldots,\infty\}$ then we have $\forall i\in\{l,\ldots,\infty\}$ that $\inf(\{x_j|j\in\{l,\ldots,\infty\}\})\leqslant x_i\in\mathbb{R}$ and $\inf(\{y_j|j\in\{l,\ldots,\infty\}\})\leqslant y_i\in\mathbb{R}$ so that $\inf(\{x_j|j\in\{l,\ldots,\infty\}\})+\inf(\{y_j|j\in\{l,\ldots,\infty\}\})$ is well defined and $\inf(\{x_j|j\in\{l,\ldots,\infty\}\})+\inf(\{y_j|j\in\{l,\ldots,\infty\}\})\leqslant x_i+y_i$ So $\inf(\{x_j|j\in\{l,\ldots,\infty\}\})+\inf(\{y_j|j\in\{l,\ldots,\infty\}\})\leqslant\inf(\{x_j+y_j|j\in\{l,\ldots,\infty\}\})$ hence we have

$$\sup\left(\{\inf(\{x_j|j\in\{l,\ldots,\infty\}\})+\inf(\{y_j|j\in\{l,\ldots,\infty\}\})|l\in\{k,\ldots,\infty\}\}\right)\leqslant\sup\left(\{\inf(\{x_j+y_j|j\in\{l,\ldots,\infty\}\})|l\in\{k,\ldots,\infty\}\}\right)=\liminf_{i\to\infty}(x_i+y_i) \tag{17.22}$$

Now as $\forall l\in\inf(\{x_j|j\in\{l,\ldots,\infty\}\})\leqslant\inf(\{x_j|j\in\{l+1,\ldots,\infty\}\})$ and $\inf(\{y_j|j\in\{l,\ldots,\infty\}\})\leqslant\inf(\{y_j|j\in\{l+1,\ldots,\infty\}\})$ we can use 17.73 giving $\sup(\{\inf(\{x_j|j\in\{l,\ldots,\infty\}\})|l\in\{k,\ldots,\infty\}\})+\sup(\{\inf(\{y_j|j\in\{l,\ldots,\infty\}\})\})=\sup(\{\inf(\{x_j|j\in\{l,\ldots,\infty\}\})+\inf(\{y_j|j\in\{l,\ldots,\infty\}\})|l\in\{k,\ldots,\infty\}\})$. Applying this together with 17.22 and the definition of $\liminf_{i\to\infty}$ we have finally

$$\liminf_{i\to\infty}x_i+\liminf_{i\to\infty}y_i\leqslant\liminf_{i\to\infty}(x_i+y_i)$$

b. Define $\{\bar{x}_i\}_{i\in\{k,\ldots,\infty\}},\{\bar{y}_i\}_{i\in\{k,\ldots,\infty\}}$ by $\bar{x}_i=-x_i$, $\bar{y}_i=-y_i$ then using (1) we have that $\liminf_{i\to\infty}\bar{x}_i+\liminf_{i\to\infty}\bar{y}_i\leqslant\liminf_{i\to\infty}(\bar{x}_i+\bar{y}_i)\Rightarrow\liminf_{i\to\infty}(-x_i)+\liminf_{i\to\infty}(-y_i)\leqslant\liminf_{i\to\infty}(-(x_i+y_i))$, so using 17.72 we have $-\limsup_{i\to\infty}x_i+\left(-\limsup_{i\to\infty}y_i\right)\leqslant-\limsup_{i\to\infty}(x_i+y_i)$. Hence

$$\limsup_{i\to\infty}(x_i+y_i)\leqslant\limsup_{i\to\infty}x_i+\limsup_{i\to\infty}y_i$$ □

The following theorem shows the relation between both types of limits.

Theorem 17.75. *Let* $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ *then* $\liminf_{i\to\infty} x_i \leqslant \limsup_{i\to\infty} x_i$

Proof. Fix $n\in\{k,\ldots,\infty\}$ then $\forall m\in\{k,\ldots,\infty\}$ we have either

$\boldsymbol{m\in\{k,\ldots,n-1\}}$**.** then $m\leqslant n$ so that $\{x_i|i\in\{n,\ldots,\infty\}\}\subseteq\{x_i|i\in\{m,\ldots,\infty\}\}$ hence $\inf(\{x_i|i\in\{m,\ldots,\infty\}\})\leqslant\inf(\{x_i|i\in\{n,\ldots,\infty\}\})\leqslant x_n\leqslant\sup(\{x_i|i\in\{n,\ldots,\infty\}\})$

$\boldsymbol{m\in\{n,\ldots,\infty\}}$**.** then $\inf(\{x_i|i\in\{m,\ldots,\infty\}\})\leqslant x_n\leqslant\sup(\{x_i|i\in\{n,\ldots,\infty\}\})$

so we have

$$\forall m\in\{k,\ldots,\infty\}\text{ we have }\inf(\{x_i|i\in\{m,\ldots,\infty\}\})\leqslant x_n\leqslant\sup(\{x_i|i\in\{n,\ldots,\infty\}\})$$

and thus

$$\sup(\{\inf(\{x_i|i\in\{m,\ldots,\infty\}\})|m\in\{k,\ldots,\infty\}\})\leqslant\sup(\{x_i|i\in\{n,\ldots,\infty\}\})$$

giving that

$$\sup(\{\inf(\{x_i|i\in\{m,\ldots,\infty\}\})|m\in\{k,\ldots,\infty\}\})\leqslant\inf(\{\sup(\{x_i|i\in\{n,\ldots,\infty\}\})|n\in\{k,\ldots,\infty\}\})$$

proving

$$\liminf_{n\to\infty} x_n\leqslant\limsup_{n\to\infty} x_n$$

proving the theorem. □

Theorem 17.76. *Let* $\{x_i\}_{i\in\{1,\ldots,n\}}$ *be a sequence of real numbers then we have*

1. *If* $\lim_{i\to\infty} x_i$ *exists (in* $\mathbb{R}$*) then* $\liminf_{i\to\infty} x_i=\limsup_{i\to\infty} x_i=\lim_{i\to\infty} x_i$
2. *If* $\liminf_{i\to\infty} x_i, \limsup_{i\to\infty} x_i\in\mathbb{R}$ *and* $\liminf_{i\to\infty} x_i=\limsup_{i\to\infty} x_i$ *then* $\lim_{i\to\infty} x_i$ *exists and* $\liminf_{i\to\infty} x_i=\limsup_{i\to\infty} x_i=\lim_{i\to\infty} x_i$

Proof.

1. Assume that $\lim_{i\to\infty} x_i$ exists and that $\lim_{i\to\infty} x_i=x$. Assume now that $\liminf_{i\to\infty} x_i<x$ then take $\varepsilon=x-\liminf_{i\to\infty} x_i$. By the existance of a limit there exists a N such that if $n\geqslant N$ we have that $|x_n-x|<\frac{\varepsilon}{2}$. As $\sup(\{\inf(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{1,\ldots,\infty\}\})=\liminf_{i\to\infty} x_i$ we have $\inf(\{x_i|i\in\{N,\ldots,\infty\}\})\leqslant\liminf_{i\to\infty} x_i<\liminf_{i\to\infty} x_i+\frac{\varepsilon}{2}$ there exists a $M\in\{N,\ldots,\infty\}$ such that $\inf(\{x_i|i\in\{N,\ldots,\infty\}\})\leqslant x_M<\liminf_{i\to\infty} x_i+\frac{\varepsilon}{2}$ then we have $x_M<x-\varepsilon+\frac{\varepsilon}{2}=x-\frac{\varepsilon}{2}\Rightarrow 0<x-x_M-\frac{\varepsilon}{2}\Rightarrow 0<\frac{\varepsilon}{2}<x-x_M=|x-x_M|<\frac{\varepsilon}{2}$ (as $N\leqslant M$) so we reach the contradiction . So we must have
$$x\leqslant\liminf_{i\to\infty} x_i \tag{17.23}$$

Assume now that $x < \limsup_{i\rightarrow\infty} x_i$ then take $\varepsilon = \limsup_{i\rightarrow\infty} x_i - x$. By the existance of a limit there exists a N such that if $n \geqslant N$ we have that $|x_n - x| < \frac{\varepsilon}{2}$. As $\limsup_{i\rightarrow\infty} x_i = \inf(\{\sup(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{1,\ldots,\infty\}\})$ we have that $\limsup_{i\rightarrow\infty} x_i - \frac{\varepsilon}{2} < \limsup_{i\rightarrow\infty} x_i \leqslant \sup(\{x_i|i\in\{N,\ldots,\infty\}\})$ and thus there exists a $M\in\{N,\ldots.,\infty\}$ such that $\limsup_{i\rightarrow\infty} x_i - \frac{\varepsilon}{2} < x_M \leqslant \sup(\{x_i|i\in\{N,\ldots,\infty\}\}) \Rightarrow \varepsilon + x - \frac{\varepsilon}{2} < x_M \Rightarrow 0 < \frac{\varepsilon}{2} < x_M - x = |x_M - x| < \frac{\varepsilon}{2}$ (as $N \leqslant M$) giving contradiction. So we must have that

$$\limsup_{i\rightarrow\infty} x_i \leqslant x \tag{17.24}$$

So using 17.75, 17.23 and 17.24 we have that $x \leqslant \liminf_{i\rightarrow\infty} x_i \leqslant \limsup_{i\rightarrow\infty} x_i \Rightarrow \liminf_{i\rightarrow\infty} x_i = \limsup_{i\rightarrow\infty} x_i = x$

2. Assume that $\liminf_{i\rightarrow\infty} x_i = \limsup_{i\rightarrow\infty} x_i = x$ then if $\varepsilon > 0$ we have $x - \varepsilon < x = \liminf_{i\rightarrow\infty} x_i = \sup(\{\inf(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{1,\ldots,\infty\}\})$ so there exists a $M\in\{1,\ldots,\infty\}$ such that $x - \varepsilon < \inf(\{x_i|i\in\{M,\ldots,\infty\}\}) \leqslant x$ and thus $\forall n \geqslant M$ we have $x - \varepsilon < x_n$ so we have

$$\exists M\in\{1,\ldots,\infty\} \text{ such that } \forall n \geqslant M \text{ we have } x - \varepsilon < x_n \tag{17.25}$$

As also $\inf(\{\sup(\{x_i|i\in\{k,\ldots,\infty\}\})|k\in\{1,\ldots,\infty\}\}) = \limsup_{i\rightarrow\infty} x_i = x < x + \varepsilon$ there exists a $N\in\{1,\ldots,\infty\}$ such that $x \leqslant \sup(\{x_i|i\in\{N,\ldots,\infty\}\}) < x + \varepsilon$ giving

$$\exists N\in\{1,\ldots,\infty\} \text{ such that } \forall n \geqslant N \text{ we have } x_n < x + \varepsilon \tag{17.26}$$

So using 17.25,17.26 we have if $n \geqslant \max(N, M)$ that $x - \varepsilon < x_n < x + \varepsilon \Rightarrow |x - x_n| < \varepsilon$ proving that $\lim_{i\rightarrow\infty} x_i = x$ □

Motivated by the above theorem we can extend the notation of a limit on $\mathbb{R}$ to $\bar{\mathbb{R}}$ as follows

Definition 17.77. *Let $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ be a sequence of extended reals then $\lim_{i\rightarrow\infty} x_i$ exists iff $\liminf_{i\rightarrow\infty} x_i = \limsup_{i\rightarrow\infty} x_i$ and we have then $\lim_{i\rightarrow\infty} x_i = \liminf_{i\rightarrow\infty} x_i = \limsup_{i\rightarrow\infty} x_i$. Note that by 17.69 the limit, if it exists, is independent of the value of k (as is expressed in the notation).*

Remark 17.78. To avoid execessive notation if we use $\lim_{i\rightarrow\infty} x_i = x$ then we autimatical assume that the limit exists.

Example 17.79. Let $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ be defined by $x_i = x \in \bar{\mathbb{R}}$ then $\lim_{i\rightarrow\infty} x_i = x$

Proof. $\sup(\{\inf(\{x_i|i\in\{n,\ldots,\infty\}\})|n\in\{k,\ldots,\infty\}\}) = \sup(\{\inf(\{x\})\}) = x = \inf(\{\sup(\{x\})\}) = \inf(\{\sup(\{x_i|i\in\{n,\ldots,\infty\}\})|n\in\{k,\ldots,\infty\}\}) \Rightarrow \lim_{i\rightarrow\infty} x_i = x$ □

The limit of a sequence of extended reals is independent of a translation of the index

Proposition 17.80. *Let $\{x_i\}_{i \in \{k, \ldots, \infty\}}$ be a sequence of extended reals and $n \in \mathbb{N}$ then if $\lim_{i \rightarrow \infty} x_i$ exists we have that $\lim_{i \rightarrow \infty} x_{i+n}$ exists and $\lim_{i \rightarrow \infty} x_i = \lim_{i \rightarrow \infty} x_{i+n}$*

Proof. As $\lim_{i \rightarrow \infty} x_i$ exists we have by definition that $\liminf_{i \rightarrow \infty} x_i = \limsup_{i \rightarrow \infty} x_i$ and $\lim_{i \rightarrow \infty} x_i = \liminf_{i \rightarrow \infty} x_i$. Hence we have $\liminf_{i \rightarrow \infty} x_{i+n} \underset{17.70}{=} \liminf_{i \rightarrow \infty} x_i = \lim_{i \rightarrow \infty} x_i = \limsup_{i \rightarrow \infty} x_i = \limsup_{i \rightarrow \infty} x_{i+n}$ □

Theorem 17.81. *Let $x \in \bar{\mathbb{R}}$, $\{x_i\}_{i \in \mathbb{N}}$ be such that $\forall i \in \mathbb{N}$ we have that $x + x_i$ is well defined and $\lim_{i \rightarrow \infty} x_i$ exists then $\lim_{i \rightarrow \infty} (x_i + x)$ exists and $\lim_{i \rightarrow \infty} (x + x_i) = x + \lim_{i \rightarrow \infty} x_i$*

Proof. By assumption we have that $\liminf_{i \rightarrow \infty} x_i = \limsup_{i \rightarrow \infty} x_i = \lim_{i \rightarrow \infty} x_i$. Now $\liminf_{i \rightarrow \infty} (x + x_i) \underset{17.71}{=} x + \liminf_{i \rightarrow \infty} x_i = x + \lim_{i \rightarrow \infty} x_i = x + \limsup_{i \rightarrow \infty} x_i \underset{17.71}{=} \left(\limsup_{i \rightarrow \infty} (x + x_i) \right)$ proving that

$$\lim_{i \rightarrow \infty} (x + x_i) = x + \lim_{i \rightarrow \infty} x_i$$

□

In the theorem 17.76 we have proved that if a sequence of reals converges in $\mathbb{R}$ then it converges also in $\bar{\mathbb{R}}$. We intend now to prove that increasing (decreasing) sequences in $\bar{\mathbb{R}}$ converges and that the limit is equal to sup (inf) as in $\mathbb{R}$. To do this we need the following lemma.

Lemma 17.82. *Let $\{x_i\}_{i \in \{n, \ldots, \infty\}}$ be a increasing (or decreasing) sequence of extended real numbers then if $k \in \{n, \ldots, \infty\}$ we have $\sup(\{x_i | i \in \{n, \ldots, \infty\}\}) = \sup(\{x_i | i \in \{k, \ldots, \infty\}\})$ (or $\inf(\{x_i | i \in \{n, \ldots, \infty\}\}) = \inf(\{x_i | i \in \{k, \ldots, \infty\}\})$)*

Proof. Let $k \in \{n, \ldots, \infty\}$

increasing sequence. As $\{x_i | i \in \{k, \ldots, \infty\}\} \subseteq \{x_i | i \in \{n, \ldots, \infty\}\}$ we must have

$$\sup(\{x_i | i \in \{k, \ldots, \infty\}\}) \leqslant \sup(\{x_i | i \in \{n, \ldots, \infty\}\}).$$

Next if $x \in \{x_i | i \in \{n, \ldots, \infty\}\}$ then $\exists m \in \{n, \ldots, \infty\}$ such that $x = x_m$ we have for m the following possibilities

$\boldsymbol{m \in \{n, \ldots, k-1\}}$**.** then $x = x_m \leqslant_{k-1 \leqslant k} x_k \in \{x_i | i \in k, \ldots, \infty\}$

$\boldsymbol{m \in \{k, \ldots, \infty\}}$**.** then $x \leqslant x_m \in \{x_i | i \in \{k, \ldots, \infty\}\}$

so that by 2.172 we have

$$\sup(\{x_i | i \in \{n, \ldots, \infty\}\}) \leqslant \sup(\{x_i | i \in \{k, \ldots, \infty\}\}).$$

decreasing sequence. As $\{x_i | i \in \{k, \ldots, \infty\}\} \subseteq \{x_i | i \in \{n, \ldots, \infty\}\}$ we must have

$$\inf(\{x_i | i \in \{n, \ldots, \infty\}\}) \leqslant \inf(\{x_i | i \in \{k, \ldots, \infty\}\}).$$

Next if $x\in\{x_i|i\in\{1,\ldots,\infty\}\}$ then $\exists m\in\{1,\ldots,\infty\}$ such that $x=x_m$ and we have the following cases for m

$m\in\{n,\ldots,k-1\}$. then $x=x_m\geqslant x_k\in\{x_i|i\in\{k,\ldots,\infty\}\}$

$m\in\{k,\ldots,\infty\}$. then $x=x_m\in\{x_i|i\in\{k,\ldots,\infty\}\}$

so by 2.172 we have

$$\inf(\{x_i|i\in\{k,\ldots,\infty\}\})\leqslant\inf(\{x_i|i\in\{n,\ldots,\infty\}\}).$$ □

Theorem 17.83. *Let $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ be a increasing (or decreasing) sequence of extended reals then $\lim_{i\to\infty} x_i$ exists and $\sup(\{x_i|i\in\{k,\ldots,\infty\}\})=\lim_{i\to\infty} x_i$ (or $\inf\left(\{x_i|i\in\{k,\ldots,\infty\}\}=\lim_{i\to\infty} x_i\right)$*

Proof.

increasing sequence. First $\limsup_{i\to\infty} x_i=\inf(\{\sup(\{x_i|i\in\{n,\ldots,\infty\}\})|n\in\{k,\ldots,\infty\}\})\underset{17.82}{=}\inf(\{\sup(\{x_i|i\in\{n,\ldots,\infty\}\})|n\in\{k,\ldots,\infty\}\})=\sup(\{x_i|i\in\{k,\ldots,\infty\}\})$. Next $\liminf_{i\to\infty} x_i=\sup(\{\inf(\{x_i|i\in\{n,\ldots,\infty\}\})|n\in\{k,\ldots,\infty\}\})\underset{\{x_i\}_{i\in\{1,\ldots,\infty\}}\text{ is increasing}}{=}\sup(\{x_k|k\in\{1,\ldots,\infty\}\})$. Proving that $\limsup_{i\to\infty} x_i=\liminf_{i\to\infty} x_i=\sup(\{x_i|i\in\{1,\ldots,\infty\}\})$.

decreasing sequence. First $\liminf_{i\to\infty} x_i=\sup(\{\inf(\{x_i|i\in\{n,\ldots,\infty\}\})|n\in\{k,\ldots,\infty\}\})\underset{17.82}{=}\sup(\{\inf(\{x_i|k\in\{1,\ldots,\infty\}\})|n\in\{k,\ldots,\infty\}\})=\inf(\{x_i|i\in\{k,\ldots,\infty\}\})$. Next $\limsup_{i\to\infty} x_i=\inf(\{\sup(\{x_i|i\in\{n,\ldots,\infty\}\})|n\in\{k,\ldots,\infty\}\})\underset{\{x_i\}_{i\in\{1,\ldots,\infty\}}\text{ is decreasing}}{=}\inf(\{x_i|i\in\{k,\ldots,\infty\}\})$. Proving that $\limsup_{i\to\infty} x_i=\liminf_{i\to\infty} x_i=\inf(\{x_i|i\in\{1,\ldots,\infty\}\})$. □

The above theorem motivates the following definition

Definition 17.84. *Let $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ be a sequence of extended reals then if we say that $x_n\uparrow x$ we mean that $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ is increasing and that $\lim_{n\to\infty} x_n=x$. Likewise $x_n\downarrow x$ means that $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ is decreasing and that $\lim_{n\to\infty} x_n=x$. Note that using 17.83 we have that $x_n\uparrow x$ is equivalent with $x=\sup(\{x_n|n\in\{k,\ldots,\infty\})$ and $x_n\downarrow x$ is equivalen with $x=\inf(\{x_n|n\in\{k,\ldots,\infty\}\})$*

For functions $f: A\to B$ where $A,B\subseteq\bar{\mathbb{R}}$ and $x\in\bar{\mathbb{R}}$ a limit point we have also a concept of a limit (see 17.48) $\lim_{x\to x_0} f(x)$. We show now the similarities between the two limits.

Theorem 17.85. *Let A be either $[1,\infty]$ or $[1,\infty[$, $B\subseteq\bar{\mathbb{R}}$ and $f: A\to B$ a increasing (decreasing function) then if we define $\{f_n\}_{n\in\{1,\ldots,\infty\}}$ by $f_n=f(n)$ we have that $\lim_{x\to\infty} f(x)$, $\lim_{n\to\infty} f_n$ exists and $\lim_{x\to\infty} f(x)=\lim_{n\to\infty} f_n$. In other words if f is a increasing (descreasing function) then $f_n\uparrow\lim_{x\to\infty} f(x)$ (or $f_n\downarrow\lim_{x\to\infty} f(x)$.*

Proof. First of all using 17.60 we have that $\lim_{x \rightarrow \infty} f(x)$ exists and is equal to $\sup(\{f(x)|\infty > x \geqslant 1\})$ if f is ascending and equal to $\inf(\{f(x)|\infty > x \geqslant 1\})$ if f is descending. Also as $i < i+1$ we have if f is increasing that $f_i = f(i) \leqslant f(i+1) = f_{i+1}$ and if f is decreasing then $f_{i+1} = f(i+1) \leqslant f(i) = f_i$. So that $\{f_n\}_{n \in \{k, \ldots, \infty\}}$ is increasing (f increasing) or $\{f_n\}_{n \in \{k, \ldots, \infty\}}$ (f decreasing). Using the previous theorem we have then that $\lim_{n \rightarrow \infty} f_n$ exists and is equal to $\sup(\{f_n|n \in \{1, \ldots, \infty\}\})$ if f is increasing and equal to $\inf(\{f_n|n \in \{1, \ldots, \infty\}\})$ if f is decreasing. We prove now equality, consider now the following cases

***f* is increasing.** As $\{f_n|n \in \{1, \ldots, \infty\}\} = \{f(n)|n \in \{1, \ldots, \infty\}\} \subseteq \{f(x)|\infty > x \geqslant 1\}$ we have that $\sup(\{f_n|n \in \{1, \ldots, n\}\}) \leqslant \sup(\{f(x)|\infty > x \geqslant 1\})$. Also if $y \in \{f(x)|\infty > x \geqslant 1\}$ then there exists a $\infty > x \geqslant 1$ such that $y = f(x)$, so there exists a $m \in \mathbb{N}_0$ such that $m \geqslant x$ and then $f_m = f(m) \geqslant f(x) = y$ proving by 2.172 that $\sup(\{f(x)|\infty > x \geqslant 1\}) \leqslant \sup(\{f_n|n \in \{1, \ldots, n\}\})$. So we have $\lim_{x \rightarrow \infty} f(x) = \sup(\{f(x)|\infty > x \geqslant 1\}) = \sup(f_n|n \in \{1, \ldots, \infty\}) = \lim_{n \rightarrow 1\infty} f_n$.

***f* is decreasing.** As $\{f_n|n \in \{1, \ldots, \infty\}\} = \{f(n)|n \in \{1, \ldots, \infty\}\} \subseteq \{f(x)|\infty > x \geqslant 1\}$ we have that $\inf(\{f_n|n \in \{1, \ldots, n\}\}) \geqslant \inf(\{f(x)|\infty > x \geqslant 1\})$. Also if $y \in \{f(x)|\infty > x \geqslant 1\}$ then there exists a $\infty > x \geqslant 1$ such that $y = f(x)$, so there exists a $m \in \mathbb{N}_0$ such that $m \geqslant x$ and then $f_m = f(m) \leqslant f(x) = y$ proving by 2.172 that $\inf(\{f(x)|\infty > x \geqslant 1\}) \geqslant \inf(\{f_n|n \in \{1, \ldots, n\}\})$. So we have $\lim_{x \rightarrow \infty} f(x) = \inf(\{f(x)|\infty > x \geqslant 1\}) = \inf(f_n|n \in \{1, \ldots, \infty\}) = \lim_{n \rightarrow \infty} f_n$. □

Example 17.86. Let $x \in \bar{\mathbb{R}}$, $k, l \in \mathbb{N}$ with $k \leqslant l$ and let $\{x_i\}_{i \in \{k, \ldots., \infty\}}$ be a sequence of extended reals such that $\forall i \geqslant l$ we have $x_i = x$ then $\lim_{i \rightarrow \infty} x_i = x$

Proof. $\liminf_{i \rightarrow \infty} x_i = \sup(\{\inf(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\}) = \sup(\{\inf(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{l, \ldots, \infty\}\}) \underset{x_i = x \text{ if } k \geqslant l}{=} \sup(\{\inf(\{x\})|n \in \{l, \ldots, \infty\}\}) = x = \inf(\{\sup(\{x\})|n \in \{l, \ldots, \infty\}\}) = \inf(\{\sup(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{l, \ldots, \infty\}\}) = \inf(\{\sup(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\}) = \limsup_{i \rightarrow \infty} x_i$. □

The idea of a limit in the reals is that the sequence approaches its limit the higher the index is (as is expressed in the ε definition), we show now a alternative definition of the limit in $\bar{\mathbb{R}}$ that embed this idea.

Theorem 17.87. *Let $\{x_i\}_{i \in \{k, \ldots, \infty\}}$ be a sequence of extended reals then we have the following equivalences*

1. $\lim_{i \rightarrow \infty} x_i = \infty$ *if and only* $\forall C \in \mathbb{R}_+$ *there exists a* $N \in \{k, \ldots, \infty\}$ *such that* $\forall n \geqslant N$ *we have* $C < x_n$
2. $\lim_{i \rightarrow \infty} x_i = -\infty$ *if and only* $\forall C \in \mathbb{R}_+$ *there exists a* $N \in \{k, \ldots, \infty\}$ *such that* $\forall n \geqslant N$ *we have* $x_n < -C$
3. $\lim_{i \rightarrow \infty} x_i = x \in \mathbb{R}$ *if and only* $\forall \varepsilon \in \mathbb{R}_+$ *there exists a* $N \in \{k, \ldots, \infty\}$ *such that* $\forall n \geqslant N$ *we have that* $x - \varepsilon < x_n < x + \varepsilon$ *(Note as* $x, \varepsilon \in \mathbb{R}$ *we have that* $x - \varepsilon < x_n < x + \varepsilon$ *is equivalent with* $x_n \in \mathbb{R} \wedge |x - x_n| < \varepsilon$*)*

Proof.

1.

- **⇒.** If $\lim_{i\to\infty} x_i = \infty$ then $\liminf_{i\to\infty} x_i = \sup(\{\inf(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\}) = \infty$, so by 17.11 we have that $\{\inf(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\}$ is not bounded above, so if $C \in \mathbb{R}_+$ then $\exists N \in \{k, \ldots, \infty\}$ such that $C < \inf(\{x_i|i \in \{N, \ldots, \infty\}\}) \Rightarrow \forall n \geqslant N$ we have $C < \inf(\{x_i|i \in \{N, \ldots, \infty\}\}) \leqslant x_n \Rightarrow C < x_n$.

- **⇐.** Assume that $\sup(\{\inf(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\}) < \infty$ then by the hypothese there exists a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $\sup(\{\inf(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\}) - 1 < x_n \Rightarrow \sup(\{\inf(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\}) < \inf(\{x_i|i \in \{N, \ldots, \infty\}\}) \Rightarrow \sup(\{\inf(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\}) < \inf(\{x_i|i \in \{N, \ldots, \infty\}\}) \leqslant \sup(\{\inf(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\})$ a contradiction. So we must have $\infty = \sup(\{\inf(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\}) = \liminf_{i\to\infty} x_i \leqslant_{(17.75)} \limsup_{i\to\infty} x_i \leqslant \infty$ proving that $\liminf_{i\to\infty} x_i = \infty = \limsup_{i\to\infty} x_i$ and thus $\lim_{i\to\infty} x_i$

2.

- **⇒.** If $\lim_{i\to\infty} x_i = -\infty$ then $\limsup_{i\to\infty} x_i = \inf(\{\sup(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\}) = -\infty$ so by 17.11 we have that $\{\sup(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\}$ is not bounded below, so if $C \in \mathbb{R}_+$ then $\exists N \in \{k, \ldots, \infty\}$ such that $\sup(\{x_i|i \in \{N, \ldots, \infty\}\}) < C$ so $\forall n \geqslant N$ we have $x_n \leqslant \sup(\{x_i|i \in \{N, \ldots, \infty\}\}) < C \Rightarrow x_n < C$

- **⇐.** Assume that $-\infty < \inf(\{\sup(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\})$ then by the hypothese there exists a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $x_n < \inf(\{\sup(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\}) + 1 \Rightarrow \sup(\{x_i|i \in \{N, \ldots, \infty\}\}) < \inf(\{\sup(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\}) \Rightarrow \inf(\{\sup(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\}) \leqslant \sup(\{x_i|i \in \{N, \ldots, \infty\}\}) < \inf(\{\sup(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\})$ a contradiction. So $-\infty = \inf(\{\sup(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\}) = \limsup_{i\to\infty} x_i \geqslant \liminf_{i\to\infty} x_i \geqslant -\infty$ proving that $\lim_{i\to\infty} x_i = -\infty$

3.

- **⇒.** If $\lim_{i\to\infty} x_i = x \in \mathbb{R}$ then $\sup(\{\inf(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\}) = x = \inf(\{\sup(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\})$ then if $\varepsilon \in \mathbb{R}_+$ we have

  a. $x - \varepsilon < \sup(\{\inf(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\}) = x$ so that $\exists N_1 \in \{k, \ldots, \infty\}$ such that $x - \varepsilon < \inf(\{x_i|i \in \{N_1, \ldots, \infty\}\}) \leqslant x \Rightarrow \forall n \geqslant N_1$ we have $x - \varepsilon < \inf(\{x_i|i \in \{N_1, \ldots, \infty\}\}) \leqslant x_n \Rightarrow x - \varepsilon \leqslant x_n$

  b. $x = \inf(\{\sup(\{x_i|i \in \{n, \ldots, \infty\}\})|n \in \{k, \ldots, \infty\}\}) < x + \varepsilon$ so that $\exists N_2 \in \{k, \ldots, \infty\}$ such that $x \leqslant \sup(\{x_i|i \in \{N_2, \ldots, \infty\}\}) < x + \varepsilon \Rightarrow \forall n \geqslant N$ we have $x_n \leqslant \sup(\{x_i|i \in \{N_2, \ldots, \infty\}\}) < x + \varepsilon \Rightarrow x_n < x + \varepsilon$

Using (a) en (b) we have if $n \geqslant \max(N_1, N_2)$ that $x - \varepsilon < x_n < x + \varepsilon \Rightarrow |x - x_n| < \varepsilon$

$\Leftarrow$. Assume that $\exists x \in \mathbb{R}$ such that $\forall \varepsilon \in \mathbb{R}_+$ there exists a $N_\varepsilon \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $x - \varepsilon < x_n < x + \varepsilon$. Given $\varepsilon \in \mathbb{R}_+$ there exists then a N_ε such that $\sup(\{x_i | i \in \{N_\varepsilon, \ldots, \infty\}\}) \leqslant x + \varepsilon$ and $x - \varepsilon \leqslant \inf(\{x_i | i \in \{N_\varepsilon, \ldots, \infty\}\})$ so we have [as $N_\varepsilon \in \{k, \ldots, \infty\}$] $\limsup_{i \to \infty} x_i = \inf(\{\sup(\{x_i | i \in \{n, \ldots, \infty\}\}) | n \in \{k, \ldots, \infty\}\}) \leqslant x + \varepsilon$ and

$x - \varepsilon \leqslant \sup(\{\inf(\{x_i | i \in \{n, \ldots, \infty\}\}) | n \in \{k, \ldots, \infty\}\}) = \liminf_{i \to \infty} x_i$ giving [using 17.75]

$$-\infty < x - \varepsilon \leqslant \liminf_{i \to \infty} x_i \leqslant \limsup_{i \to \infty} x_i \leqslant x + \varepsilon < \infty \tag{17.27}$$

Assume now that $\liminf_{i \to \infty} x_i < x$, take then $\varepsilon = \frac{x - \liminf_{i \to \infty} x_i}{2} > 0$ so by 17.27 we have then that $x - \left(\frac{x - \liminf_{i \to \infty} x_i}{2}\right) \leqslant \liminf_{i \to \infty} x_i \Rightarrow \frac{x}{2} + \frac{\liminf_{i \to \infty} x_i}{2} \leqslant \liminf_{i \to \infty} x_i \Rightarrow \frac{x}{2} \leqslant \frac{\liminf_{i \to \infty} x_i}{2} < \frac{x}{2}$ a contradiction. So we must have

$$x \leqslant \liminf_{i \to \circ 2} x_i \tag{17.28}$$

Assume that $x < \limsup_{i \to \infty} x_i$, take then $\varepsilon = \frac{\limsup_{i \to \infty} x_i - x}{2} > 0$ so by 17.27 we have that $\limsup_{i \to \infty} x_i \leqslant x + \frac{\limsup_{i \to \infty} x_i - x}{2} = \frac{x}{2} + \frac{\limsup_{i \to \infty} x_i}{2} \Rightarrow \frac{\limsup_{i \to \infty} x_i}{2} \leqslant \frac{x}{2} \Rightarrow \frac{x}{2} < \frac{x}{2}$ a contradiction. So we must have

$$\limsup_{i \to \infty} x_i \leqslant x \tag{17.29}$$

Using 17.27, 17.28 and 17.29 we have then that $\liminf_{i \to \infty} x_i = x = \limsup_{i \to \infty} x_i$ proving that

$$\lim_{i \to \infty} x_i = x$$

□

Corollary 17.88. *Let $\{x_i\}_{i \in \{k, \ldots, \infty\}}$ be a sequence of extended reals with $\lim_{i \to \infty} x_i = x$ then $\lim_{i \to \infty} |x_i| = |x|$*

Proof. We have three cases to consider for x

$\boldsymbol{x \in \mathbb{R}}$. using the previous theorem (see 17.87) we have $\forall \varepsilon \in \mathbb{R}_+$ that $\exists N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $x_n \in \mathbb{R} \wedge |x - x_n| < \varepsilon \underset{||x| - |x_n|| < \varepsilon}{\Rightarrow} x_n \in \mathbb{R} \wedge ||x| - |x_n|| < \varepsilon$ proving by the previous theorem again that $\lim_{i \to \infty} |x_i| = |x|$.

$\boldsymbol{x = \infty}$. using the previous theorem (see 17.87) we have $\forall C \in \mathbb{R}_+$ that there exists a $N \in \{, \ldots., \infty\}$ such that $\forall n \geqslant N$ we have $C \leqslant x_n \underset{0 \leqslant C \Rightarrow 0 \leqslant x_n}{\Rightarrow} C \leqslant |x_n|$ proving by the previous theorem that $\lim_{i \to \infty} x_n = \infty = |x|$

$x = -\infty$. orem (see 17.87) we have $\forall C \in \mathbb{R}_+$ that there exists a $N \in \{, \ldots., \infty\}$ such that $\forall n \geqslant N$ we have $x_n \leqslant -C \Rightarrow C \leqslant -x_n \underset{0 \leqslant C \Rightarrow 0 \leqslant -x_n \Rightarrow x_n \leqslant 0}{\Rightarrow} C \leqslant |x_n|$ proving by the previous theorem that $\lim_{i \rightarrow \infty} x_n = \infty = |x|$ □

We show now that the limit in $\bar{\mathbb{R}}$ has similar properties as the limit in $\mathbb{R}$

Theorem 17.89. *The limit in $\bar{\mathbb{R}}$ has the following properties*

1. *Let $A \in \bar{\mathbb{R}}$, $\{x_i\}_{i \in \{k, \ldots, \infty\}}$ a sequence of extended reals with a limit so that $\forall i \in \{k, \ldots, \infty\}$ we have $x_i \leqslant A$ [or $A \leqslant x_i$] then $\lim_{i \rightarrow \infty} x_i \leqslant A$ [or $A \leqslant \lim_{i \rightarrow \infty} x_i$]*
2. *Let $\{x_i\}_{i \in \{k, \ldots, \infty\}}$, $\{y_i\}_{i \in \{k, \ldots, \infty\}}$ be sequences of extended reals with a limit such that $\forall i \in \{k, \ldots, \infty\}$ we have $x_i \leqslant y_i$ then $\lim_{i \rightarrow \infty} x_i \leqslant \lim_{i \rightarrow \infty} y_i$*
3. *If $\{x_i\}_{i \in \{k, \ldots, \infty\}}$ has a limit and $\alpha \in \mathbb{R}$ then $\{\alpha \cdot x_i\}_{i \in \{k, \ldots, \infty\}}$ has a limit and $\lim_{i \rightarrow \infty} (\alpha \cdot x_i) = \alpha \cdot \lim_{i \rightarrow \infty} x_i$*
4. *Let $\{x_i\}_{i \in \{k, \ldots, \infty\}} \subseteq \bar{\mathbb{R}}$ and $\{y_i\}_{i \in \{k, \ldots, \infty\}} \subseteq \bar{\mathbb{R}}$ be sequences of extended reals with limits such that $\left(\lim_{i \rightarrow \infty} x_i\right) + \left(\lim_{i \rightarrow \infty} y_i\right)$ is well defined and there exists a $n \in \{k, \ldots, \infty\}$ such that $\forall i \in \{n, \ldots, \infty\}$ we have that $x_i + y_i$ is well defined then for $\{x_i + y_i\}_{i \in \{n, \ldots, \infty\}}$ we have $\lim_{i \rightarrow \infty} (x_i + y_i) = \left(\lim_{i \rightarrow \infty} x_i\right) + \left(\lim_{i \rightarrow \infty} y_i\right)$*
5. *Let $\{x_i\}_{i \in \{k, \ldots, \infty\}} \subseteq \bar{\mathbb{R}}$ and $\{y_i\}_{i \in \{k, \ldots, \infty\}} \subseteq \bar{\mathbb{R}}$ be sequences of extended reals with limits such that for $x = \lim_{i \rightarrow \infty} x_i$ and $y = \lim_{i \rightarrow \infty}$ we have if $x = 0$ then $y \neq \infty$, $-\infty$ and if $y = 0$ then $x \neq \infty, -\infty$ then $\lim_{i \rightarrow \infty} (x_i \cdot y_i) = \left(\lim_{i \rightarrow \infty} x_i\right) \cdot \left(\lim_{i \rightarrow \infty} y_i\right)$*
6. *Let $\{x_i\}_{i \in \{k, \ldots., \infty\}}, \{y_i\}_{i \in \{k, \ldots, \infty\}}$ be sequences of extended reals with limits then we have*

   a. $\lim_{i \rightarrow \infty} (\min(x_i, y_i)) = \min\left(\lim_{i \rightarrow \infty} x_i, \lim_{i \rightarrow \infty} y_i\right)$

   b. $\lim_{i \rightarrow \infty} (\max(x_i, y_i)) = \max\left(\lim_{i \rightarrow \infty} x_i, \lim_{i \rightarrow \infty} y_i\right)$

Proof.

1.

   a. If $\forall i \in \{k, \ldots, \infty\}$ we have $x_i \leqslant A$ then $\sup(\{x_i | i \in \{k, \ldots, \infty\}\}) \leqslant A \Rightarrow \forall n \in \{k, \ldots, \infty\}$ we have $\sup(\{x_i | i \in \{n, \ldots, \infty\}\}) \leqslant \sup(\{x_i | i \in \{k, \ldots, \infty\}\}) \leqslant A$ and thus $\inf(\{\sup(\{x_i | i \in \{n, \ldots, \infty\}\}) | n \in \{k, \ldots, \infty\}\}) \leqslant A \underset{\text{definition of limit}}{\Rightarrow} \lim_{i \rightarrow \infty} x_i \leqslant A$

   b. If $\forall i \in \{k, \ldots, \infty\}$ we have $A \leqslant x_i$ then $A \leqslant \inf(\{x_i | i \in \{k, \ldots, \infty\}\}) \Rightarrow \forall n \in \{k, \ldots, \infty\}$ we have $A \leqslant \inf(\{x_i | i \in \{k, \ldots, \infty\}\}) \leqslant \inf(\{x_i | i \in \{n, \ldots, \infty\}\})$ so that $A \leqslant \sup(\{\inf(\{x_i | i \in \{n, \ldots, \infty\}\}) | n \in \{k, \ldots, \infty\}\}) \underset{\text{defnition of limit}}{\Rightarrow} A \leqslant \lim_{i \rightarrow \infty} x_i$

2. As $\forall i \in \{k, \ldots, \infty\}$ we have $x_i \leqslant y_i$, we have $\forall n \in \{k, \ldots, n\}$ that $\forall i \in \{n, \ldots, \infty\}$ we have $x_i \leqslant y_i \underset{2.172}{\Rightarrow} \sup(\{x_i | i \in \{n, \ldots, \infty\}\}) \leqslant \sup(\{y_i | i \in \{n, \ldots, \infty\}\}) \underset{2.172}{\Rightarrow} \inf(\{\sup(\{x_i | i \in \{n, \ldots, \infty\}\}) | n \in \{k, \ldots, \infty\}\}) \leqslant \inf(\{\sup(\{y_i | i \in \{n, \ldots, \infty\}\}) | n \in \{k, \ldots, \infty\}\}) \underset{\text{definition of limit}}{\Rightarrow} \lim_{i \rightarrow \infty} x_i \leqslant \lim_{i \rightarrow \infty} y_i$
3. Consider the following possible cases for $\alpha \in \mathbb{R}$

$\alpha = 0$. then $\forall i \in \{k, \ldots, \infty\}$ we have $\alpha \cdot x_i = 0$ so that $\forall n \in \{k, \ldots, \infty\}$ we have $\sup(\{\alpha \cdot x_i | i \in \{n, \ldots, \infty\}\}) = \inf(\{\alpha \cdot x_i | i \in \{n, \ldots, \infty\}\}) = 0$ so that $\inf(\{\sup(\{\alpha \cdot x_i | i \in \{n, \ldots, \infty\}\}) | n \in \{k, \ldots, \infty\}\}) = 0 = \sup(\{\inf(\{\alpha \cdot x_i | i \in \{n, \ldots, \infty\}\}) | n \in \{k, \ldots, \infty\}\}) \Rightarrow \lim_{i \rightarrow \infty} (\alpha \cdot x_i) = 0 = \alpha \cdot \lim_{i \rightarrow \infty} x_i$ proving the theorem in this case.

$\alpha \in \mathbb{R} \setminus \{0\}$. then we have the following possibilities for α

$\alpha > 0$. consider now the following possibilities for $\lim_{i \rightarrow \infty} x_i$

$\lim_{i \rightarrow \infty} x_i = \infty$. Let now $C \in \mathbb{R}_+$ then there exists a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $x_n > \frac{C}{\alpha} \Rightarrow \alpha \cdot x_n > C$ proving that $\lim_{i \rightarrow \infty} (\alpha \cdot x_i) = \infty = \alpha \cdot \infty$ and thus $\lim_{i \rightarrow \infty} (\alpha \cdot x_i) = \alpha \cdot \lim_{i \rightarrow \infty} x_i$

$\lim_{i \rightarrow \infty} x_i = -\infty$. Let now $C \in \mathbb{R}_+$ then there exists a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $x_n < -\frac{C}{\alpha} \Rightarrow \alpha \cdot x_n < -C \Rightarrow \lim_{i \rightarrow \infty} (\alpha \cdot x_i) = -\infty = \alpha \cdot (-\infty)$ and thus $\lim_{i \rightarrow \infty} (\alpha \cdot x_i) = \alpha \cdot \lim_{i \rightarrow \infty} x_i$

$\lim_{i \rightarrow \infty} x_i \in \mathbb{R}$. Take $x = \lim_{i \rightarrow \infty} x_i$, let now $\varepsilon \in \mathbb{R}_+$ then there exists a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $|x - x_n| < \frac{\varepsilon}{\alpha} \Rightarrow x - \frac{\varepsilon}{\alpha} < x_n < x - \frac{\varepsilon}{\alpha} \Rightarrow \alpha \cdot x - \varepsilon < \alpha \cdot x_n < \alpha \cdot x + \varepsilon \Rightarrow |\alpha \cdot x - \alpha \cdot x_n| < \varepsilon$ proving that $\lim_{i \rightarrow \infty} (\alpha \cdot x_i) = \alpha \cdot x = \alpha \cdot \lim_{i \rightarrow \infty} x_i$

$\alpha < 0$. consider now the following possibilities for $\lim_{i \rightarrow \infty} x_i$

$\lim_{i \rightarrow \infty} x_i = \infty$. Let now $C \in \mathbb{R}_+$ then there exists a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $x_n > \frac{C}{-\alpha} \Rightarrow \alpha \cdot x_n < -C \Rightarrow \lim_{i \rightarrow \infty} (\alpha \cdot x_i) = -\infty = \alpha \cdot \infty$ proving that $\lim_{i \rightarrow \infty} (\alpha \cdot x_i) = \alpha \cdot \lim_{i \rightarrow \infty} x_i$

$\lim_{i \rightarrow \infty} x_i = -\infty$. Let now $C \in \mathbb{R}_+$ then there exists a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $x_n < \frac{-C}{-\alpha} \Rightarrow C < \alpha \cdot x_n \Rightarrow \lim_{i \rightarrow \infty} x_i = \infty = \alpha \cdot (-\infty)$ proving that $\lim_{i \rightarrow \infty} (\alpha \cdot x_i) = \alpha \cdot \lim_{i \rightarrow \infty} x_i$

$\lim_{i\rightarrow\infty} x_i \in \mathbb{R}$. Take $x = \lim_{i\rightarrow\infty} x_i$ and let $\varepsilon > 0$ then there exists a $N \in \{k, ..., \infty\}$ such that $|x - x_n| < \frac{\varepsilon}{-\alpha} \Rightarrow x + \frac{\varepsilon}{\alpha} < x_n < x - \frac{\varepsilon}{\alpha} \Rightarrow \alpha \cdot x - \varepsilon < x_n < \alpha \cdot x + \varepsilon \Rightarrow |\alpha \cdot x - \alpha \cdot x_n| < \varepsilon$ proving that $\lim_{i\rightarrow\infty} (\alpha \cdot x_i) = \alpha \cdot x = \alpha \cdot \lim_{i\rightarrow\infty} x_i$

So in all cases we have that $\lim_{i\rightarrow\infty} (\alpha \cdot x_i) = \alpha \cdot \lim_{i\rightarrow\infty} x_i$

4. We have that $\left(\lim_{i\rightarrow\infty} x_i\right) + \left(\lim_{i\rightarrow\infty} y_i\right)$ is only well defined in the following cases

$\lim_{i\rightarrow\infty} x_i = \infty = \lim_{i\rightarrow\infty} y_i$. then there exists $N_1, N_2 \in \{k, ..., \infty\}$ such that $\forall i \geqslant N_1$ we have $1 < x_i$ and $\forall i \geqslant N_2$ we have $1 < y_i$ so if $n = \max(N_1, N_2)$ we have $\forall i \geqslant n$ that $x_i + y_i$ is well defined. Let now $C \in \mathbb{R}_+$ then there exists $M_1, M_2 \in \{k, ..., \infty\}$ such that $\forall i \geqslant M_1$ we have $x_i > \frac{C}{2}$ and $\forall i \geqslant M_2$ we have $y_i > \frac{C}{2}$ so that if $n \geqslant \max(n, M_1, M_2)$ we have $x_i + y_i > \frac{C}{2} + \frac{C}{2} = C$ proving that for $\{x_i + y_i\}_{i \in \{n, ..., \infty\}}$ we have $\lim_{i\rightarrow\infty} (x_i + y_i) = \infty = \infty + \infty = \left(\lim_{i\rightarrow\infty} x_i\right) + \left(\lim_{i\rightarrow\infty} y_i\right)$

$\lim_{i\rightarrow\infty} x_i = -\infty = \lim_{i\rightarrow\infty} y_i$. then there exists $N_1, N_2 \in \{k, ..., \infty\}$ such that $\forall i \geqslant N_1$ we have $x_i < -1$ and $\forall i \geqslant N_2$ we have $y_i < -1$ so if $n = \max(N_1, N_2)$ we have $\forall i \geqslant n$ that $x_i + y_i$ is well defined. Let now $C \in \mathbb{R}_+$ then there exists $M_1, M_2 \in \{k, ..., \infty\}$ such that $\forall i \geqslant M_1$ we have $x_i < \frac{-C}{2}$ and $\forall i \geqslant M_2$ we have $y_i < \frac{-C}{2}$ so that if $n \geqslant \max(n, M_1, M_2)$ we have $x_i + y_i < \frac{-C}{2} + \frac{-C}{2} = C$ proving that for $\{x_i + y_i\}_{i \in \{n, ..., \infty\}}$ we have $\lim_{i\rightarrow\infty} (x_i + y_i) = -\infty = -\infty + (-\infty) = \left(\lim_{i\rightarrow\infty} x_i\right) + \left(\lim_{i\rightarrow\infty} y_i\right)$

$\lim_{i\rightarrow\infty} x_i = \infty \wedge \lim_{i\rightarrow\infty} y_i \in \mathbb{R}$. Take $y = \lim_{i-.\infty} y_i$ then there exists a $N_1 \in \{k, ..., \infty\}$ such that $\forall i \geqslant N_1$ we have $|y - y_i| < 1 \Rightarrow -\infty < y - 1 < y_i$. Also there exists a $N_2 \in \{k, ..., \infty\}$ such that $\forall i \geqslant N_1$ we have $x_i > 1 > -\infty$. So if $n = \max(N_1, N_2)$ then we have $\forall i \geqslant n$ that $x_i + y_i$ is well defined and $y_i > y - 1$. Take now $C \in \mathbb{R}_+$ then there exists a $M \in \{k, ..., n\}$ such that $\forall i \geqslant M$ we have $x_i > \max(C - (y - 1), 1)$ then if $i \geqslant N = \max(n, M)$ we have that $x_i > C - (y - 1)$ and $y_i > y - 1$ so that $x_i + y_i > C$ proving that $\lim_{i\rightarrow\infty} (x_i + y_i) = \infty = \infty + y = \left(\lim_{i\rightarrow\infty} x_i\right) + \left(\lim_{i\rightarrow\infty} y_i\right)$

$\lim_{i\rightarrow\infty} x_i = -\infty \wedge \lim_{i\rightarrow\infty} y_i \in \mathbb{R}$. Take $y = \lim_{i-.\infty} y_i$ then there exists a $N_1 \in \{k, ..., \infty\}$ such that $\forall i \geqslant N_1$ we have $|y - y_i| < 1 \Rightarrow y_i < y + 1$. Also there exists a $N_2 \in \{k, ..., \infty\}$ such that $\forall i \geqslant N_1$ we have $x_i < -1 < \infty$. So if $n = \max(N_1, N_2)$ then we have $\forall i \geqslant n$ that $x_i + y_i$ is well defined and $y_i < y + 1$. Take now $C \in \mathbb{R}_+$ then there exists a $M \in \{k, ..., n\}$ such that $\forall i \geqslant M$ we have $x_i < \min(-C - (y + 1), -1)$ then if $i \geqslant N = \max(n, M)$ we have that $x_i < -C - (y + 1)$ and $y_i < y + 1$ so that $x_i + y_i < -C$ proving that $\lim_{i\rightarrow\infty} (x_i + y_i) = -\infty = -\infty + y = \left(\lim_{i\rightarrow\infty} x_i\right) + \left(\lim_{i\rightarrow\infty} y_i\right)$

$\boldsymbol{\lim_{i\to\infty} x_i \in \mathbb{R} \wedge \lim_{i\to\infty} y_i = \infty.}$ This is equivalent with the case $\lim_{i\to\infty} x_i = \infty \wedge \lim_{i\to\infty} y_i \in \mathbb{R}$ if we interchage $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ and $\{y_i\}_{i\in\{k,\ldots,\infty\}}$

$\boldsymbol{\lim_{i\to\infty} x_i \in \mathbb{R} \wedge \lim_{i\to\infty} y_i = -\infty.}$ This is equivalent with the case $\lim_{i\to\infty} x_i = -\infty \wedge \lim_{i\to\infty} y_i \in \mathbb{R}$ if we interchage $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ and $\{y_i\}_{i\in\{k,\ldots,\infty\}}$

$\boldsymbol{\lim_{i\to\infty} x_i \in \mathbb{R} \wedge \lim_{i\to\infty} y_i \in \mathbb{R}.}$ Let $\lim_{i\to\infty} x_i = x$ and $\lim_{i\to\infty} y_i = y$, then there exists a $N_1, N_2 \in \{k, \ldots, \infty\}$ such that if $i \geqslant N_1$ we have $|x_i - x| < 1 \Rightarrow x - 1 < x_i < x + 1 \Rightarrow x_i \in \mathbb{R}$ and if $i \geqslant N_2$ we have $|y - y_i| < 1 \Rightarrow y - 1 < y_i < y + 1 \Rightarrow y_i \in \mathbb{R}$. Take $n = \max(N_1, N_2)$ then if $i \geqslant n$ we have that $x_i + y_i$ is well defined. Let now $\varepsilon \in \mathbb{R}_+$ then there exists $M_1, M_2 \in \{k, \ldots, \infty\}$ such that $\forall i \geqslant M_1$ we have $|x - x_i| < \frac{\varepsilon}{2} \Rightarrow x - \frac{\varepsilon}{2} < x_i < x + \frac{\varepsilon}{2}$ and $\forall i \geqslant M_2$ we have $|y - y_i| < \frac{\varepsilon}{2} \Rightarrow y - \frac{\varepsilon}{2} < y_i < y + \frac{\varepsilon}{2}$. So if $i \geqslant \max(M_1, M_2, n)$ we have $\left(x - \frac{\varepsilon}{2}\right) + \left(y - \frac{\varepsilon}{2}\right) < x_i + y_i < \left(x + \frac{\varepsilon}{2}\right) + \left(y + \frac{\varepsilon}{2}\right) \Rightarrow (x + y) - \varepsilon < x_i + y_i < (x + y) + \varepsilon \Rightarrow |(x + y) - (x_i + y_i)| < \varepsilon$ proving that $\lim_{i\to\infty}(x_i + y_i) = x + y = \left(\lim_{i\to\infty} x_i\right) + \left(\lim_{i\to\infty} y_i\right)$

So in all cases we have that $\exists n \in \{k, \ldots, \infty\}$ such that $\forall i \geqslant n$ we have $x_i + y_i$ is well defined, and for $\{x_i + y_i\}_{i\in\{k,\ldots,\infty\}}$ we have $\lim_{i\to\infty}(x_i + y_i) = \left(\lim_{i\to\infty} x_i\right) + \left(\lim_{i\to\infty} y_i\right)$.

5. Let $x = \lim_{i\to\infty} x_i$ and $y = \lim_{i\to\infty}$ we have if $x = 0$ then $y \neq \infty, -\infty$ and if $y = 0$ then $x \neq \infty, -\infty$ then we must consider the following remaining cases:

$\boldsymbol{x = \infty \wedge y = \infty.}$ Let $C \in \mathbb{R}_+$ then there exists $N_1, N_2 \in \{k, \ldots, \infty\}$ such that $\forall i \geqslant N_1$, $\forall j \geqslant N_2$ we have $1 \leqslant x_i$ and $C \leqslant y_i$. Take $n = \max(N_1, N_2)$ then if $i \geqslant n$ we have $1 \leqslant x_i \wedge C \leqslant y_i \Rightarrow C \leqslant C \cdot x_i \wedge C \cdot x_i \leqslant x_i \cdot y_i \Rightarrow C \leqslant x_i \cdot y_i$. This proves that $\lim_{i\to\infty}(x_i \cdot y_i) = \infty = \infty \cdot \infty = \left(\lim_{i\to\infty} x_i\right) \cdot \left(\lim_{i\to\infty} y_i\right)$

$\boldsymbol{x = \infty \wedge y = -\infty.}$ Let $C \in \mathbb{R}_+$ then there exists $N_1, N_2 \in \{k, \ldots, \infty\}$ such that $\forall i \geqslant N_1$, $\forall j \geqslant N_2$ we have $1 \leqslant x_i$ and $y_i \leqslant -C$. Take $n = \max(N_1, N_2)$ then if $i \geqslant n$ we have $1 \leqslant x_i \wedge y_i \leqslant -C \Rightarrow C \leqslant C \cdot x_i \wedge y_i \cdot x_i \leqslant -C \cdot x_i \Rightarrow -C \cdot x_i \leqslant -C \wedge y_i \cdot x_i \leqslant -C \cdot x_i \Rightarrow x_i \cdot y_i \leqslant -C$. This proves that $\lim_{i\to\infty}(x_i \cdot y_i) = -\infty = \infty \cdot -\infty = \left(\lim_{i\to\infty} x_i\right) \cdot \left(\lim_{i\to\infty} y_i\right)$

$\boldsymbol{x = \infty \wedge y \in \mathbb{R} \setminus \{0\}.}$ We have then for y two possiibilities:

$\boldsymbol{0 < y.}$ Let $C \in \mathbb{R}_+$ take $\varepsilon = \frac{y}{2} > 0$ then $\exists N_1, N_2 \in \{k, \ldots, \infty\}$ such that $\forall i \geqslant N_1, \forall j \geqslant N_2$ we have $y - \varepsilon < y_i < y + \varepsilon \wedge \frac{2 \cdot C}{y} \leqslant x_i$. Take $n = \max(N_1, N_2)$ then if $n \leqslant i$ we have $\frac{y}{2} < y_i \wedge \frac{2 \cdot C}{y} \leqslant x_i \underset{0 < \frac{2 \cdot C}{y} \wedge 0 < \frac{y}{2} < y_i}{\Rightarrow} \left(\frac{2 \cdot C}{y}\right) \cdot \frac{y}{2} \leqslant \frac{2 \cdot C}{2} y_i \wedge \frac{2 \cdot C}{y} \cdot y_i \leqslant x_i \cdot y_i \Rightarrow C \leqslant \frac{2 \cdot C}{y} \cdot y_i \wedge \frac{2 \cdot C}{y} \cdot y_i \leqslant x_i \cdot y_i \Rightarrow C \leqslant x_i \cdot y_i$. So we have $\lim_{i\to\infty}(x_i \cdot y_i) = \infty = x \cdot \infty = \left(\lim_{i\to\infty} x_i\right) \cdot \left(\lim_{i\to\infty} y_i\right)$

$y<0$. Take $\{y'_i\}_{i\in\{k,\ldots,\infty\}}$ defined by $y'_i=-y_i$ then $\lim_{i\to\infty} y'_i = \lim_{i\to\infty}(-y_i)=-y>0$ so by the previous case we have that $-\lim_{i\to\infty}(x_i\cdot y_i)=\lim_{i\to\infty}(x_i\cdot(-y_i))=\left(\lim_{i\to\infty}x_i\right)\cdot\left(\lim_{i\to\infty}(-y_i)\right)=-\left(\left(\lim_{i\to\infty}x_i\right)\cdot\left(\lim_{i\to\infty}y_i\right)\right)$ proving that $\lim_{i\to\infty}(x_i\cdot y_i)=\left(\lim_{i\to\infty}x_i\right)\cdot\left(\lim_{i\to\infty}y_i\right)$

$x=-\infty\wedge y=\infty$. this is the same as $x=\infty\wedge y=-\infty$ with x and y exchanged.

$x=-\infty\wedge y=-\infty$. Define $\{x'_i\}_{i\in\{k,\ldots,\infty\}}$, $\{y'_i\}_{i\in\{k,\ldots,\infty\}}$ by $x'_i=-x_i\wedge y'_i=-y_i$ then we have $\lim_{i\to\infty}x'_i\underset{(3)}{=}-\lim_{i\to\infty}x_i=-x>0\wedge\lim_{i\to\infty}y'_i\underset{(3)}{=}-\lim_{i\to\infty}y_i=-y>0$. As we have already proved this case we have $\lim_{i\to\infty}(x_i\cdot y_i)=\lim_{i\to\infty}(-x_i\cdot(-y_i))=\left(\lim_{i\to\infty}-x_i\right)\cdot\left(\lim_{i\to\infty}(-y_i)\right)=\left(\lim_{i\to\infty}x_i\right)\cdot\left(\lim_{i\to\infty}y_i\right)$ proving that $\lim_{i\to\infty}(x_i\cdot y_i)=\left(\lim_{i\to\infty}x_i\right)\cdot\left(\lim_{i\to\infty}y_i\right)$

$x=-\infty\wedge y\in\mathbb{R}\setminus\{0\}$. As $\lim_{i\to\infty}(-x_i)\underset{(3)}{=}\infty$ we can use the case $x=\infty\wedge y\in\mathbb{R}\setminus\{0\}$ giving $\lim_{i\to\infty}(-x_i\cdot y_i)=\left(\lim_{i\to\infty}(-x_i)\right)\cdot\lim_{i\to\infty}y_i\Rightarrow-\left(\lim_{i\to\infty}(x_i\cdot y_i)\right)=-\left(\left(\lim_{i\to\infty}x_i\right)\cdot\left(\lim_{i\to\infty}y_i\right)\right)\Rightarrow\lim_{i\to\infty}(x_i\cdot y_i)=\left(\lim_{i\to\infty}x_i\right)\cdot\left(\lim_{i\to\infty}y_i\right)$

$x\in\mathbb{R}\setminus\{0\}\wedge y=\infty$. this is the same as $x=\infty\wedge y\in\mathbb{R}\setminus\{0\}$ with x,y interchanged.

$x\in\mathbb{R}\setminus\{0\}\wedge y=-\infty$. this is the same as $x=-\infty\wedge y\in\mathbb{R}\setminus\{0\}$ with x,y interchanged.

$x\in\mathbb{R}\wedge y\in\mathbb{R}$. Take $\varepsilon>0$. First there exists a $N_1\in\{k,\ldots,\infty\}$ such that $\forall i\geqslant N_1$ we have $|x_i-x|<1\Rightarrow|x_i|\leqslant|x_i-x|+|x|<1+|x|$. Find then $N_2\in\{k,\ldots,\infty\}$ such that $\forall i\geqslant N_2$ we have $|x_i-x|<\frac{\varepsilon}{2\cdot(1+|y|)}$ and a $N_3\in\{k,\ldots,\infty\}$ such that $|y_i-y_i|<\frac{\varepsilon}{2\cdot(1+|x|)}$. Take now $n=\max(N_1,N_2)$ then if $i\geqslant n$ we have tht $|x_i\cdot y_i-x\cdot y|=|x_i\cdot y_i-x_i\cdot y+x_i\cdot y-x\cdot y|=|x_i\cdot(y_i-y)+(x_i-x)\cdot y|\leqslant|x_i|\cdot|y_i-y|+|y|\cdot|x_i-x|<(1+|x|)\cdot\frac{\varepsilon}{2\cdot(1+|x|)}+\frac{\varepsilon}{2\cdot(1+|y|)}\cdot|y|=\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$ proving that $\lim_{i\to\infty}(x_i\cdot y_i)=x\cdot y=\left(\lim_{i\to\infty}x_i\right)\cdot\left(\lim_{i\to\infty}y_i\right)$

6. Let $x=\lim_{i\to\infty}x_i$ and $y=\lim_{i\to\infty}y_i$ then we have the following cases to consider for x,y

   $x=\infty\wedge y=\infty$. then given $C\in\mathbb{R}_+$ there exists $N_1,N_2\in\{k,\ldots,\infty\}$ such that $\forall i\geqslant N_1$,$\forall j\geqslant N_2$ we have $C\leqslant x_i,C\leqslant y_j$. Take $N=\max(N_1,N_2)$ then $\forall i\geqslant N$ we have $C\leqslant\max(x_i,y_i)$ and $C\leqslant\min(x_i,y_i)$ proving that

   a. $\lim_{i\to\infty}(\min(x_i,y_i))=\infty=\min(\infty,\infty)=\min\left(\lim_{i\to\infty}x_i,\lim_{i\to\infty}y_i\right)$

b. $\lim_{i\rightarrow\infty}(\max(x_i, y_i)) = \infty = \max(\infty, \infty) = \max\left(\lim_{i\rightarrow\infty} x_i, \lim_{i\rightarrow\infty} y_i\right)$

$\boldsymbol{x = \infty \wedge y = -\infty}$**.** then given $C \in \mathbb{R}_+$ there exists $N_1, N_2 \in \{k, \ldots, \infty\}$ such that $\forall i \geqslant N_1$,$\forall j \geqslant N_2$ we have $C \leqslant x_i, y_j \leqslant -C \underset{-C \leqslant C}{\Rightarrow} y_j \leqslant -C \leqslant C \leqslant x_i$. Take $N = \max(N_1, N_2)$ then $\forall i \in \mathbb{N}_0$ we have $C \leqslant \max(x_i, y_i)$ and $\min(x_i, y_i) \leqslant -C$ proving that

a. $\lim_{i\rightarrow\infty}(\min(x_i, y_i)) = -\infty = \min(\infty, -\infty) = \min\left(\lim_{i\rightarrow\infty} x_i, \lim_{i\rightarrow\infty} y_i\right)$

b. $\lim_{i\rightarrow\infty}(\max(x_i, y_i)) = \infty = \max(\infty, -\infty) = \max\left(\lim_{i\rightarrow\infty} x_i, \lim_{i\rightarrow\infty} y_i\right)$

$\boldsymbol{x = \infty \wedge y \in \mathbb{R}}$**.** then given $C \in \mathbb{R}_+$ and $\varepsilon > 0$ there exists a $N_1, N_2 \in \{k, \ldots, \infty\}$ such that $\forall i \geqslant N_1$, $\forall j \geqslant N_2$ we have $\max(C, |y| + \varepsilon) \leqslant x_i \wedge |y - y_i| < \varepsilon \Rightarrow C \leqslant x_i \wedge |y - y_i| < \varepsilon \wedge y_i < y + \varepsilon \leqslant |y| + \varepsilon \leqslant x_i$. If we take then $N = \max(N_1, N_2)$ we have $\forall i \geqslant N$ that

a. $|\min(x_i, y_i) - x| = |y_i - x| < \varepsilon$ proving that $\lim_{i\rightarrow\infty}(\min(x_i, y_i)) = y = \min(\infty, y) = \min\left(\lim_{i\rightarrow\infty} x_i, \lim_{i\rightarrow\infty} y_i\right)$

b. $C \leqslant x_i = \max(x_i, y_i)$ proving that $\lim_{i\rightarrow\infty}(\max(x_i, y_i)) = \infty = \max(\infty, y) = \max\left(\lim_{i\rightarrow\infty} x_i, \lim_{i\rightarrow\infty} y_i\right)$

$\boldsymbol{x = -\infty \wedge y = \infty}$**.** this reduces to the case $x = \infty \wedge y = -\infty$ if we interchange x and y

$\boldsymbol{x = -\infty \wedge y = -\infty}$**.** then given $C \in \mathbb{R}_+$ there exists $N_1, N_2 \in \{k, \ldots, \infty\}$ such that $\forall i \geqslant N_1$,$\forall j \geqslant N_2$ we have $x_i \leqslant -C$, $y_j \leqslant -C$. Take $N = \max(N_1, N_2)$ then $\forall i \geqslant N$ we have $\max(x_i, y_i) \leqslant -C$ and $\min(x_i, y_i) \leqslant -C$ proving that

a. $\lim_{i\rightarrow\infty}(\min(x_i, y_i)) = -\infty = \min(-\infty, -\infty) = \min\left(\lim_{i\rightarrow\infty} x_i, \lim_{i\rightarrow\infty} y_i\right)$

b. $\lim_{i\rightarrow\infty}(\max(x_i, y_i)) = -\infty = \max(-\infty, -\infty) = \max\left(\lim_{i\rightarrow\infty} x_i, \lim_{i\rightarrow\infty} y_i\right)$

$\boldsymbol{x = -\infty \wedge y \in \mathbb{R}}$**.**

$\boldsymbol{x \in \mathbb{R} \wedge y = \infty}$**.** this reduce to the case $x = \infty \wedge y \in \mathbb{R}$ if we interchange x and y

$\boldsymbol{x \in \mathbb{R} \wedge y = -\infty}$**.** this reduce to the case $x = -\infty \wedge y \in \mathbb{R}$ if we interchange x and y

$x\in\mathbb{R}\wedge y\in\mathbb{R}$. then given $\varepsilon>0$ there exists $N_1,N_2\in\{k,\ldots,\infty\}$ such that $\forall i\geqslant N_1,\forall j\geqslant N_2$ we have $|x_i-x|<\varepsilon\wedge|y_j-y|<\varepsilon\Leftrightarrow x-\varepsilon<x_i<x+\varepsilon\wedge y-\varepsilon<y_j<y+\varepsilon$. Take then $N=\max(N_1,N_2)$ then we have $\forall i\geqslant N$ that

a. $\min(x,y)-\varepsilon=\min(x-\varepsilon,y-\varepsilon)<x_i,y_i\Rightarrow\min(x,y)-\varepsilon<\min(x_i,y_i)$ and $\min(x_i,y_i)<x+\varepsilon,y+\varepsilon\Rightarrow\min(x_i,y_i)<\min(x+\varepsilon,y+\varepsilon)<\min(x,y)+\varepsilon$ proving that $|\min(x_i,y_i)-\min(x,y)|<\varepsilon$ giving $\lim_{i\rightarrow\infty}(\min(x_i,y_i))=\min(x,y)=\min\left(\lim_{i\rightarrow\infty}x_i,\lim_{i\rightarrow\infty}y_i\right)$

b. $x-\varepsilon,y-\varepsilon<\max(x_i,y_i)\Rightarrow\max(x,y)-\varepsilon=\max(x-\varepsilon,y-\varepsilon)<\max(x_i,y_i)$ and $x_i,y_i<\max(x+\varepsilon,y+\varepsilon)=\max(x,y)+\varepsilon$ proving that $|\max(x_i,y_i)-\max(x,y)|<\varepsilon$ giving $\lim_{i\rightarrow\infty}(\max(x_i,y_i))=\max(x,y)=\max\left(\lim_{i\rightarrow\infty}x_i,\lim_{i\rightarrow\infty}y_i\right)$ □

Note 17.90. The condition that α in the above theorem is real is required as the following example shows. The sequence $\left\{\frac{(-1)^i}{i}\right\}_{i\in\{1,\ldots,\infty\}}$ converges to 0 but $\left\{\infty\cdot\frac{(-1)^i}{i}\right\}_i=\{(-1)^i\cdot\infty\}_{i\in\{1,\ldots,\infty\}}$ does not converges.

Note 17.91. The extra conditions in (5) are needed, for example take $\{x_i\}_{i\in\mathbb{N}}$ defined by $x_i=\frac{1}{i}$ and $\{y_i\}_{i\in\mathbb{N}}$ defined by $y_i=i^2$ then $\lim_{i\rightarrow\infty}x_i=0$ and $\lim_{i\rightarrow\infty}y_i=\infty$ but $\lim_{i\rightarrow\infty}(x_i\cdot y_i)=\lim_{i\rightarrow\infty}i=\infty\neq\infty\cdot 0=0$

To prove that $\lim_{i\rightarrow\infty}\left(\frac{1}{x_i}\right)=\frac{1}{\lim_{i\rightarrow\infty}x_i}$ we first need a lemma

Lemma 17.92. *Let $\{x_i\}_{i\in\{k,\ldots,\infty\}}\subseteq\mathbb{R}$ be a sequence of real numbers such that $\forall i\in\{k,\ldots,\infty\}$ we have $x_i\neq 0$ and $0\neq\lim_{i\rightarrow\infty}x_i\in\mathbb{R}$ then $\inf(\{|x_i||i\in\{k,\ldots,\infty\}\})>0$*

Proof. Take $x=\lim_{i\rightarrow\infty}x_i$ and define $\varepsilon=\frac{|x|}{2}>0$, as $x\in\mathbb{R}$ there exists a $N\in\{k+1,\ldots,\infty\}$ such that $\forall i\geqslant N$ we have $|x_i-x|<\varepsilon$. Assume that $\exists i\geqslant N$ such that $|x_i|<\varepsilon$ then we have $|x|=|x-x_i+x_i|\leqslant|x-x_i|+|x_i|<\varepsilon+\varepsilon=2\cdot\varepsilon=|x|\Rightarrow|x|<|x|$ a contradiction. So we must have

$$\forall i\geqslant N\vDash|x_i|\geqslant\varepsilon$$

If we take also $M'=\min(\{|x_i||i\in\{k,\ldots,N\}\})>0$ [as $\forall i\in\{k,\ldots,\infty\}$ we have $x_i\neq 0$) and $M=\min(M',\varepsilon)>0$ then $\forall i\in\{k,\ldots,\infty\}$ we have $|x_i|\geqslant M>0$ so that $\inf(\{|x_i||i\in\{k,\ldots,\infty\}\})\geqslant M>0$ as we must prove. □

Theorem 17.93. *Let $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ be a sequence of extended reals such that $\lim_{i\rightarrow\infty}x_i\neq 0$ then there exists a $n\in\{k,\ldots,\infty\}$ such that $\forall i\in\{h,\ldots,\infty\}$ we have $x_i\neq 0$ and for $\{x_i\}_{i\in\{n,\ldots,n\}}$ we have that $\lim_{i\rightarrow\infty}\frac{1}{x_i}=\frac{1}{\lim_{i\rightarrow\infty}x_i}$*

Proof. Let $x = \lim_{i \rightarrow \infty} x_i$ then we must consider the following cases for x

$\boldsymbol{x = \infty}$**.** First there exists a $n \in \{k, \ldots, \infty\}$ such that $\forall i \in \{k, \ldots, \infty\}$ we have $1 \leqslant x_i \Rightarrow \forall i \in \{m, \ldots, \infty\}$ we have $x_i \neq 0$. Take now $\varepsilon > 0$ then by the Archimedean property of the reals (see 9.55) there exists a $C \in \mathbb{N}$ such that $0 < \frac{1}{C} < \varepsilon$. As $x = \infty$ there exists a $N \in \{n, \ldots, \infty\}$ such that $\forall i \geqslant N$ we have $x_i \geqslant C > 0$ hence $0 < \frac{1}{x_i} < \frac{1}{C} < \varepsilon \Rightarrow \left|\frac{1}{x_i} - 0\right| = \frac{1}{x_i} < \varepsilon \Rightarrow$ proving that $\lim_{i \rightarrow \infty} \left(\frac{1}{x_i}\right) = 0 = \frac{1}{\infty} = \frac{1}{\lim_{i \rightarrow \infty} x_i}$-0

$\boldsymbol{x = -\infty}$**.** First there exists a $n \in \{k, \ldots, \infty\}$ such that $\forall i \in \{k, \ldots, \infty\}$ we have $x_i \leqslant -1 \Rightarrow \forall i \in \{m, \ldots, \infty\}$ we have $x_i \neq 0$. Take now $\varepsilon > 0$ then by the Archimedean property of the reals (see 9.55) there exists a $C \in \mathbb{N}$ such that $0 < \frac{1}{C} < \varepsilon \Rightarrow -\varepsilon < \frac{1}{-C} < 0$. As $x = -\infty$ there exists a $N \in \{n, \ldots, \infty\}$ such that $\forall i \geqslant N$ we have $x_i \leqslant -C < 0 \Rightarrow 0 < C \leqslant -x_i$ hence $0 < \frac{1}{-x} \leqslant \frac{1}{C} < \varepsilon \Rightarrow -\varepsilon < \frac{1}{x_i} < 0 \Rightarrow \frac{1}{x_i} \in]0 - \varepsilon, 0 + \varepsilon[$ proving that $\lim_{i \rightarrow \infty} \left(\frac{1}{x_i}\right) = 0 = \frac{1}{-\infty} = \frac{1}{\lim_{i \rightarrow \infty} x_i}$

$\boldsymbol{x \in \mathbb{R} \setminus \{0\}}$**.** As $x \neq 0$ we have that $0 < |x|$ take $\delta = \frac{|x|}{2} > 0$ then there exists a $n \in \{k, \ldots, \infty\}$ such that $\forall i \geqslant n$ we have $|x - x_i| < \delta = \frac{|x|}{2}$, if now there exists a $i \geqslant n$ such that $x_i = 0$ then we have $|x| < \frac{|x|}{2}$ a contradiction and if $|x_i| = \infty$ we have $\infty < \frac{|x|}{2} \in \mathbb{R}$ also a contradiction.. So $\forall i \in \{n, \ldots, \infty\}$ we have $0 \neq x_i \in \mathbb{R}$. Using the above lemma 17.92 we have that $m = \inf(\{|x_i| | i \in \{n, \ldots, \infty\}\}) > 0$ hence $\forall i \in \{n, \ldots, \infty\}$ we have that $x_i \geqslant m > 0$. Take now $\varepsilon > 0$ then there exists a $N \in \{n, \ldots, \infty\}$ such that $\forall i \geqslant N$ we have $|x - x_i| < \varepsilon \cdot |x| \cdot m$, then $\left|\frac{1}{x_i} - \frac{1}{x}\right| = \left|\frac{x_i - x}{x_i \cdot x}\right| = \frac{|x_i - x|}{|x_i| \cdot |x|} \leqslant \frac{|x_i - x|}{m \cdot |x|} < \frac{\varepsilon \cdot |x| \cdot m}{|x| \cdot m} = \varepsilon$ proving that $\lim_{i \rightarrow \infty} \left(\frac{1}{x_i}\right) = \frac{1}{\lim_{i \rightarrow \infty} x_i}$ □

Example 17.94. We have that $\lim_{n \rightarrow \infty} n = \infty$ and $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$

Proof. As $\{n\}_{n \in \mathbb{N}}$ is increasing $\lim_{n \rightarrow \infty} n = \sup(\{n | n \in \mathbb{N}\}) = \sup(\mathbb{N})$ which exists. Assume that $\sup(\mathbb{N}) < \infty$ then there exists a $n \in \mathbb{N}$ such that $\sup(\mathbb{N}_{0_n}) < n \leqslant \sup(\mathbb{N}) < n$ leading to the contradiction $n < n$ hence we have $\lim_{n \rightarrow \infty} n = \infty$. Using the previous theorem 17.93 we have then $\lim_{n \rightarrow \infty} \left(\frac{1}{n}\right) = \frac{1}{\lim_{n \rightarrow \infty} n} = \frac{1}{\infty} = 0$. □

Theorem 17.95. *Let $\{x_i\}_{i \in \{k, \ldots, \infty\}} \subseteq \bar{\mathbb{R}}$ with $\lim_{i \rightarrow \infty} x_i = x \in \mathbb{R}$ then $\lim_{i \rightarrow \infty} |x - x_i| = 0$. Further if $x \in \mathbb{R}$ is such that $\lim_{i \rightarrow \infty} |x - x_i| = 0$ then $\lim_{i \rightarrow \infty} x_i = x$*

Proof. First as $x \in \mathbb{R}$ we have that $\forall i \in \mathbb{N}$ $x - x_i$ is well defined. Also by 17.87 we have that $\forall \varepsilon \in \mathbb{R}_+$ there exists a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $|x_n - x| < \varepsilon \Rightarrow ||x_n - x| - 0| = |x_n - x| < \varepsilon$ proving by 17.87 again that $\lim_{i \rightarrow \infty} |x_i - x| = 0$. If $\lim_{i \rightarrow \infty} |x - x_i| = 0 \in \mathbb{R}$ then by 17.87 we have that $\forall \varepsilon \in \mathbb{R}_+$ there $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $||x_n - x| - 0| < \varepsilon \Rightarrow |x_n - x| < \varepsilon$ proving by 17.87 again that $\lim_{i \rightarrow \infty} x_i = x = 0$ □

Note 17.96. The condition that $x \in \mathbb{R}$ in the above definition is needed. For example take $\{x_i\}_{i \in \mathbb{N}}$ defined by $x_i = i$ then $\lim_{i \to \infty} x_i = \infty$ however $\lim_{i \to \infty} |x_i - x| = \lim_{i \to \infty} |x_i - \infty| = \infty \neq 0$

17.2.3 Series in $\bar{\mathbb{R}}_+$

To avoid any problems with non defined sums in $\bar{\mathbb{R}}$ (like $-\infty + \infty$ and $+\infty + (-\infty)$ we restrict ourself to $\bar{\mathbb{R}}_+ = \{x \in \bar{\mathbb{R}} | x \geqslant 0\}$ so that by 17.26 $\langle \bar{\mathbb{R}}_+, + \rangle$ is a Abelian semi-group. All the theorems, definitions and propositions for Abelian semi groups in section 10.1 will then apply. In particular we can use the definition of a finite sum in 10.2 to define $\sum_{i=0}^{n} x_i$ where $\forall i \in \{0, \ldots, n\}$ we have $x_i \in \bar{\mathbb{R}}_+$. So using 10.5 we have the following:

Theorem 17.97. *Let $\{x_i\}_{i \in \{0, \ldots, n\}}$ be a finite family of non negative extended reals [$\forall i \in \{0, \ldots, n\} \quad x_i \in \bar{\mathbb{R}}_+$] then* $\sum_{i=0}^{n} x_i = \begin{cases} x_0 \text{ if } n = 0 \\ \sum_{i=0}^{n-1} x_i + x_n \text{ if } n > 0 \end{cases}$

Using 10.21 and 10.23 we have also in general

Theorem 17.98. *Let $\{x_i\}_{i \in \{n, \ldots, m\}}$ a finite family of non negative extended reals then* $\sum_{i=n}^{m} x_i \underset{\text{def}}{=} \sum_{i=0}^{m-n} x_{i+n} = \begin{cases} x_n \text{ if } m = n \\ \sum_{i=n}^{m-1} x_i + x_m \text{ if } m > n \end{cases}$

Using the above we prove that the sum of a finite family of non negative extended reals is non negative as expressed in the following theorem.

Theorem 17.99. *Let $\{x_i\}_{i \in \{k, \ldots, n\}}$ be a finite family of non negative extended reals then $0 \leqslant \sum_{k=1}^{n} x_i$ and more general if $\{x_i\}_{i \in I}$ is a finite family of non negative extended reals then $0 \leqslant \sum_{i \in I} x_i$ (see 10.37)*

Proof. We start by induction so let $\mathcal{S} = \{n \in \{k, \ldots, \infty\} |$if $\{x_i\}_{i \in \{k, \ldots, n\}}$ is a family of non negative extended reals then $0 \leqslant \sum_{i=k}^{n} x_i\}$ then we have

- $\boldsymbol{k \in \mathcal{S}}$**.** As $0 \leqslant x_k = \sum_{i=k}^{k} x_i$
- $\boldsymbol{n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}}$**.** Let $\{x_i\}_{i \in \{k, \ldots, n+1\}}$ is a family of non negative extended reals then $\sum_{k=1}^{n+1} x_i = 0 \leqslant (\sum_{k=1}^{n} x_i) + x_{n+1}$ [as $n \in \mathcal{S}$ we have $0 \leqslant \sum_{i=k}^{n} x_i$] proving $n + 1 \in \mathcal{S}$

For the last part of the theorem if I is a finite family we have a bijection σ: $\{1, \ldots, n\} \to I$ and thus $\sum_{i \in I} x_i = \sum_{i=1}^{n} x_{\sigma(i)} \geqslant 0$ □

As a application of the above theorem we prove that series are increeasing based on there domain.

Theorem 17.100. *$\{x_i\}_{i \in I}$ is a finite family of extended non negative reals then $\forall \emptyset \neq J \subseteq I$ we have $\sum_{i \in J} x_i \leqslant \sum_{i \in I} x_i$. So if $\{x_i\}_{i \in \{1, \ldots, n\}}$ is a finite family of non negative extended reals then $\forall m \in \{1, \ldots, n\}$ we have $\sum_{i=1}^{m} x_i \leqslant \sum_{i=1}^{n} x_i$*

Proof. As $I = (I \setminus J) \bigcup J$ and $(I \setminus J) \bigcap I = \emptyset$ we have by 10.46 that $\sum_{i \in I} x_i = \sum_{i \in (I \setminus J)} x_i + \sum_{i \in J} x_i \underset{0 \leqslant \sum_{i \in I \setminus J} x_i \text{ and } 17.22}{\Rightarrow} \sum_{i \in J} x_i \leqslant \sum_{i \in I} x_i$ □

Theorem 17.101. *Let $\{x_i\}_{i \in \{n, \ldots, m\}}$ be a finite family of non negative reals then if $\alpha \in \mathbb{R}$ we have $\sum_{i=n}^{m} (\alpha \cdot x_i) = \alpha \cdot \sum_{i=n}^{m} x_i$*

Proof. We prove this by induction so let $S = \{m \in \{n, \ldots, \infty\} |$If $\{x_i\}_{i \in \{n, \ldots, m\}}$ is a family of extended reals then $\sum_{i=n}^{m} (\alpha \cdot x_i) = \alpha \cdot \sum_{i=n}^{m} x_i\}$ then we have

- $\boldsymbol{m \in S}$**.** If $\{x_i\}_{i \in \{m, \ldots, m\}}$ is a finite family of extended reals then $\sum_{i=m}^{m} (\alpha \cdot x_i) = \alpha \cdot x_m = \alpha \cdot (\sum_{i=m}^{m} x_i)$ proving that $m \in S$
- $\boldsymbol{m \in S \Rightarrow m + 1 \in S}$**.** If $\{x_i\}_{i \in n, \ldots, m+1}$ is a finite family of extended reals then $\sum_{i=n}^{m+1} (\alpha \cdot x_i) = (\sum_{i=n}^{m} (\alpha \cdot x_i)) + \alpha \cdot x_{m+1} \underset{m \in S}{=} \alpha \cdot (\sum_{i=1}^{m} x_i) + \alpha \cdot x_{m+1} \underset{17.28}{=} \alpha \cdot (\sum_{i=1}^{m+1} x_i)$ proving that $m + 1 \in S$ □

Another application is proving that the sum of limits is the limit of sums as is expressed in the following theorem.

Theorem 17.102. *If $\{\{x_{i,j}\}_{j \in \{k, \ldots, \infty\}}\}_{i \in \{l, \ldots, n\}}$ is a finite family of sequences of positive extended reals such that $\forall i \in \{l, \ldots, n\}$ we have that $\{x_{i,j}\}_{j \in \{k, \ldots, \infty\}}$ has a limit then $\{\sum_{i=l}^{n} x_{i,j}\}_{j \in \{k, \ldots, n\}}$ is well defined and $\lim_{j \to \infty} (\sum_{i=l}^{n} x_{i,j}) = \sum_{i=l}^{n} \left(\lim_{j \to \infty} x_{i,j} \right)$*

Proof. We use mathematical induction to prove this, so let $\mathcal{S} = \Big\{ n \in \{l, \ldots, \infty\} |$if $\{\{x_{i,j}\}_{j \in \{k, \ldots, \infty\}}\}_{i \in \{l, \ldots, n\}}$ is such that $\forall i \in \{l, \ldots, n\} \vDash \{x_{i,j}\}_{j \in \{k, \ldots, \infty\}}$ has a limit then $\lim_{j \to \infty} (\sum_{i=l}^{n} x_{i,j}) = \sum_{i=1}^{n} \left(\lim_{j \to \infty} x_{i,j} \right) \Big\}$, we have then :

- $\boldsymbol{l \in \mathcal{S}}$**.** If $\{\{x_{i,j}\}_{j \in \{k, \ldots, \infty\}}\}_{i \in \{l, \ldots, l\}}$ is such that $\forall i \in \{l, \ldots, l\} \vDash \{x_{i,j}\}_{j \in \{k, \ldots, \infty\}}$ has a limit then $\{\sum_{i=l}^{l} x_{i,j}\}_{j \in \{k, \ldots, \infty\}} = \{x_{l,j}\}_{j \in \{k, \ldots, \infty\}}$ has a a limit and $\lim_{j \to \infty} (\sum_{i=l}^{1} x_{i,j}) = \lim_{j \to \infty} x_{l,j} = \sum_{i=l}^{l} \left(\lim_{j \to \infty} x_{i,j} \right)$ proving that $l \in S$
- $\boldsymbol{n \in S \Rightarrow n + 1 \in \mathcal{S}}$**.** If $\{\{x_{i,j}\}_{j \in \{k, \ldots, \infty\}}\}_{i \in \{l, \ldots, n+1\}}$ is such that $\forall i \in \{l, \ldots, n + 1\} \vDash \{x_{i,j}\}_{j \in \{k, \ldots, \infty\}}$ has a limit, then as $n \in \mathcal{S}$ we have that $\{\sum_{l=1}^{n} x_{i,j}\}_{j \in \{k, \ldots, \infty\}}$ has a limit. Then $\{\sum_{i=l}^{n+1} x_{i,j}\}_{j \in \{k, \ldots, \infty\}} = \{(\sum_{i=l}^{n} x_{i,j}) + x_{n+1,j}\}_{j \in \{k, \ldots, \infty\}}$ has a limit (using 17.89 (4)) and $\lim_{j \to \infty} (\sum_{i=l}^{n+1} x_{i,j}) = \lim_{j \to \infty} ((\sum_{i=l}^{n} x_{i,j}) + x_{n+1,j}) \underset{17.89\ (4)}{=} \left(\lim_{j \to \infty} \sum_{i=l}^{n} x_{i,j} \right) + \lim_{j \to \infty} x_{n+1,j} \underset{n \in \mathcal{S}}{=} \sum_{i=l}^{n} \left(\lim_{j \to \infty} x_{i,j} \right) + \lim_{j \to \infty} x_{n+1,j} = \sum_{i=l}^{n+1} \left(\lim_{j \to \infty} x_{i,j} \right)$ proving that $n + 1 \in \mathcal{S}$ □

We have a similar theorem for limits of functions

Definition 17.103. *If $A, B \subseteq \bar{\mathbb{R}}$ $f: A \to B$, $g: A \to B$ functions then the graph $(f + g) \subseteq A \times \bar{\mathbb{R}}$ is defined by $(f + g)(x) = f(x) + g(x)$ (assuming that the sum is well defined (so no $\infty + (-\infty)$))*

Definition 17.104. *Let A, B sets $\{f_i: A \rightarrow B\}_{i \in \{l, \ldots, n\}}$ a finite family of functions then the graph $\sum_{i=l}^{n} f_i \subseteq A \times \bar{\mathbb{R}}$ is defined by*

1. $\sum_{i=l}^{l} f_i = f_l$
2. $\sum_{i=l}^{n+1} f_i = (\sum_{i=l}^{n} f_i) + f_{n+1}$ *(assuming that the sum is well defined)*

A alternative definition is then expressed in the following lemma

Lemma 17.105. *et A, B sets $\{f_i: A \rightarrow B\}_{i \in \{l, \ldots, n\}}$ a finite family of functions then the graph $\sum_{i=l}^{n} f_i \subseteq A \times \bar{\mathbb{R}}$ is defined by $(\sum_{i=l}^{n} f_i)(x) = \sum_{i=l}^{n} f_i(x)$*

Proof. The trivial proof is by induction so let $\mathcal{S} = \{n \in \mathbb{N} | \text{If} \{f_i: A \rightarrow B\}_{i \in \{l, \ldots, n\}}$ is a finite family of functions then $(\sum_{i=l}^{n} f_i)(x) = \sum_{i=l}^{n} f_i(x)\}$ then we have

$l \in \mathcal{S}$. This follows as $(\sum_{i=l}^{l} f_i)(x) = f_l(x) = \sum_{i=l}^{l} f_i(x)$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$. Let $n \in \mathcal{S}$ and take $\{f_i: A \rightarrow B\}_{i \in \{l, \ldots, n+1\}}$ a finite family of functions then $(\sum_{i=l}^{n+1} f_i)(x) = (\sum_{i=l}^{n} f_i + f_{n+1})(x) = (\sum_{i=l}^{n} f_i)(x) + f_{n+1}(x) \underset{n \in \mathcal{S}}{=} \sum_{i=l}^{n} f_i(x) + f_{n+1}(x) = \sum_{i=l}^{n+1} f_i(x)$ □

Theorem 17.106. *Let $A, B \subseteq \bar{\mathbb{R}}$, $x_0 \in \bar{\mathbb{R}}$ and $\{f_i: A \rightarrow B\}_{i \in \{1, \ldots, n\}}$ a finite family of functions then if $(\sum_{i=l}^{n} f_i)(A) \subseteq C$ we have for $\sum_{i=l}^{n} f_i: A \rightarrow C$*

1. *If $\forall i \in \{l, \ldots, n\}$ $\lim_{x \downarrow x_0} f_i(x)$ exists then $\lim_{x \downarrow x_0} (\sum_{i=l}^{n} f_i)(x) = \sum_{i=l}^{n} \lim_{x \downarrow x_0} f_i(x)$ (assuming that the sum is well defined)*
2. *If $\forall i \in \{l, \ldots, n\}$ $\lim_{x \uparrow x_0} f_i(x)$ exists then $\lim_{x \uparrow x_0} (\sum_{i=l}^{n} f_i)(x) = \sum_{i=l}^{n} \lim_{x \uparrow x_0} f_i(x)$ (assuming that the sum is well defined)*
3. *If $\forall i \in \{l, \ldots, n\}$ $\lim_{x \rightarrow x_0} f_i(x)$ exists then $\lim_{x \rightarrow x_0} (\sum_{i=l}^{n} f_i)(x) = \sum_{i=l}^{n} \lim_{x \rightarrow x_0} f_i(x)$ (assuming that the sum is well defined)*

Proof. This is proved by induction

1. Let $\mathcal{S} = \left\{n \in \{l, \ldots, \infty\} | \text{for} \{f_i: A \rightarrow B\}_{i \in \{l, \ldots, n\}} \text{ we have } \lim_{x \downarrow x_0} (\sum_{i=l}^{n} f)(x) = \sum_{i=l}^{n} \lim_{x \downarrow x_0} f(x)\right\}$ then we have

   $l \in \mathcal{S}$. If we have $\{f_i | A \rightarrow B|\}_{i \in \{l, \ldots, l\}}$ then we have $\lim_{x \downarrow x_0} (\sum_{i=l}^{l} f_i)(x) = \lim_{x \downarrow x_0} f_l(x) = \sum_{i=l}^{l} \lim_{x \downarrow x_0} f_i(x)$ proving that $l \in \mathcal{S}$

   $n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$. Take $\{f_i: A \rightarrow B\}_{i \in \{l, \ldots, n+1\}}$ then we have that $\lim_{x \downarrow x_0} (\sum_{i=l}^{n+1} f_i)(x) = \lim_{x \downarrow x_0} (\sum_{i=l}^{n} f_i + f_{n+1})(x) \underset{17.51}{=} \lim_{x \downarrow x_0} (\sum_{i=l}^{n} f_i)(x) + \lim_{x \downarrow x_0} f_{n+1}(x) \underset{n \in \mathcal{S}}{=} \sum_{i=l}^{n} \lim_{x \downarrow x_0} f_i(x) + \lim_{x \downarrow x_0} f_{n+1}(x) = \sum_{i=l}^{n+1} \lim_{x \downarrow x_0} f_i(x)$ proving that $n+1 \in \mathcal{S}$

2. Let $\mathcal{S} = \left\{ n \in \{l, \ldots, \infty\} | \text{for } \{f_i: A \to B\}_{i \in \{l, \ldots, n\}} \text{ we have } \lim_{x \uparrow x_0} (\sum_{i=l}^n f)(x) = \sum_{i=l}^n \lim_{x \uparrow x_0} f(x) \right\}$ then we have

   $l \in \mathcal{S}$. If we have $\{f_i | A \to B|\}_{i \in \{l, \ldots, l\}}$ then we have $\lim_{x \uparrow x_0} (\sum_{i=l}^l f_i)(x) = \lim_{x \uparrow x_0} f_l(x) = \sum_{i=l}^l \lim_{x \uparrow x_0} f_i(x)$ proving that $l \in \mathcal{S}$

   $n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$. Take $\{f_i: A \to B\}_{i \in \{l, \ldots, n+1\}}$ then we have that $\lim_{x \uparrow x_0} (\sum_{i=l}^{n+1} f_i)(x) = \lim_{x \uparrow x_0} (\sum_{i=l}^n f_i + f_{n+1})(x) \underset{17.51}{=} \lim_{x \uparrow x_0} (\sum_{i=l}^n f_i)(x) + \lim_{x \uparrow x_0} f_{n+1}(x) \underset{n \in \mathcal{S}}{=} \sum_{i=l}^n \lim_{x \uparrow x_0} f_i(x) + \lim_{x \uparrow x_0} f_{n+1}(x) = \sum_{i=l}^{n+1} \lim_{x \uparrow x_0} f_i(x)$ proving that $n + 1 \in \mathcal{S}$

3. Let $\mathcal{S} = \left\{ n \in \{l, \ldots, \infty\} | \text{for } \{f_i: A \to B\}_{i \in \{l, \ldots, n\}} \text{ we have } \lim_{x \to x_0} (\sum_{i=l}^n f)(x) = \sum_{i=l}^n \lim_{x \to x_0} f(x) \right\}$ then we have

   $l \in \mathcal{S}$. If we have $\{f_i | A \to B|\}_{i \in \{l, \ldots, l\}}$ then we have $\lim_{x \to x_0} (\sum_{i=l}^l f_i)(x) = \lim_{x \to x_0} f_l(x) = \sum_{i=l}^l \lim_{x \to x_0} f_i(x)$ proving that $l \in \mathcal{S}$

   $n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$. Take $\{f_i A \to B\}_{i \in \{l, \ldots, n+1\}}$ then we have that $\lim_{x \to x_0} (\sum_{i=l}^{n+1} f_i)(x) = \lim_{x \to x_0} (\sum_{i=l}^n f_i + f_{n+1})(x) \underset{17.51}{=} \lim_{x \to x_0} (\sum_{i=l}^n f_i)(x) + \lim_{x \to x_0} f_{n+1}(x) \underset{n \in \mathcal{S}}{=} \sum_{i=l}^n \lim_{x \to x_0} f_i(x) + \lim_{x \to x_0} f_{n+1}(x) = \sum_{i=l}^{n+1} \lim_{x \to x_0} f_i(x)$ proving that $n + 1 \in \mathcal{S}$ $\square$

Definition 17.107. *Let $A, B \subseteq \bar{\mathbb{R}}$, $f: A \to B$ a function , $\alpha \in \bar{\mathbb{R}}$ then the graph $\alpha \cdot f \subseteq A \times \bar{\mathbb{R}}$ is defined by $(\alpha \cdot f)(x) = \alpha \cdot f(x)$*

Corollary 17.108. *Let $A, B \subseteq \bar{\mathbb{R}}$, $x_0 \in \bar{\mathbb{R}}$ and $\{f_i: A \to B\}_{i \in \{k, \ldots, n\}}$ a finite family of functions, $\{\alpha_i\}_{i \in \{k, \ldots, n\}}$ a finite family of extended reals then if $(\sum_{i=k}^n \alpha_i \cdot f_i)(A) \subseteq C$ we have for $\sum_{i=k}^n \alpha \cdot f_i: A \to C$ that*

1. *If $\forall i \in \{k, \ldots, n\}$ $\lim_{x \downarrow x_0} f_i(x)$ exists then $\lim_{x \downarrow x_0} (\sum_{k=1}^n \alpha \cdot f_i)(x) = \sum_{i=k}^n \alpha_i \cdot \lim_{x \downarrow x_0} f_i(x)$ (assuming that the sum is well defined)*
2. *If $\forall i \in \{k, \ldots, n\}$ $\lim_{x \uparrow x_0} f_i(x)$ exists then $\lim_{x \uparrow x_0} (\sum_{i=k}^n \alpha_i \cdot f_i)(x) = \sum_{i=k}^n \alpha_i \cdot \lim_{x \uparrow x_0} f_i(x)$ (assuming that the sum is well defined)*
3. *If $\forall i \in \{k, \ldots, n\}$ $\lim_{x \to x_0} f_i(x)$ exists then $\lim_{x \to x_0} (\sum_{i=k}^n f_i)(x) = \sum_{i=k}^n \alpha_i \cdot \lim_{x \to x_0} f_i(x)$ (assuming that the sum is well defined)*

Proof. This follows from applying 17.51 on the scalar product of a scalar and a function and the previous theorem (see 17.106) $\square$

We prove now that if $\infty \notin \{x_i | i \in \{n, \ldots, m\}\}$ then $\sum_{i=n}^m x_i$ is finite and otherwise ∞

Theorem 17.109. *Let $\{x_i\}_{i \in \{n, \ldots, m\}}$ be a finite family of extended non negative reals then we have*

1. *If $\infty \in \{x_i | i \in \{n, \ldots, m\}\}$ then $\sum_{i=n}^m x_i = \infty$*

2. *If* $\infty \notin \{x_i|i \in \{n, \ldots, m\}\}$ *then* $\sum_{i=n}^{m} x_i < \infty$

Proof.

1. We prove this by induction so let $\mathcal{S} = \{m \in \{n, \ldots, \infty\}|$If $\{x_i\}_{i \in \{n, \ldots m\}}$ is *a* family in $\bar{\mathbb{R}}_+$ with $\infty \in \{x_n, \ldots, x_m\}$ then $\sum_{i=m}^{m} x_i = \infty\}$ then we have

   $n \in \mathcal{S}$. then $\infty \in \{x_n, \ldots, x_n\} \Rightarrow x_n = \infty$ and $\sum_{i=n}^{n} x_i = x_n = \infty$ proving that $n \in \mathcal{S}$

   $m \in \mathcal{S} \Rightarrow m + 1 \in \mathcal{S}$. let $\infty \in \{x_n, \ldots, x_{m+1}\}$ then we have either

   $x_{m+1} = \infty$. then $\sum_{i=n}^{m+1} x_i = \sum_{i=n}^{m} x_i + x_m = \sum_{i=n}^{m} x_i + \infty = \infty$

   $x_{m+1} \in \{x_n, \ldots, x_m\}$. then $\sum_{i=n}^{m+1} x_i = \sum_{i=n}^{m} x_i + x_{m+1} \underset{n \in \mathcal{S}}{=} \infty + x_{m+1} = \infty$

   so in all cases we have $\sum_{i=n}^{m+1} x_i = \infty$ proving that $m + 1 \in \mathcal{S}$

   Using mathematical induction we have then that $\forall m \in \{n, \ldots, \infty\}$ we have $\sum_{i=n}^{m} x_i = \infty$

2. We prove this also by induction so let $\mathcal{S} = \{m \in \{n, \ldots, \infty\}|$If $\{x_i\}_{i \in \{n, \ldots m\}}$ is *a* family in $\bar{\mathbb{R}}_+$ with $\infty \notin \{x_n, \ldots, x_m\}$ then $\sum_{i=m}^{m} x_i < \infty\}$ then we have

   $n \in \mathcal{S}$. then as $\infty \notin \{x_n, \ldots, x_n\} \Rightarrow x_n < \infty$ and thus $\sum_{i=n}^{n} x_i = x_n < \infty$ proving that $n \in \mathcal{S}$

   $m \in \mathcal{S} \Rightarrow m + 1 \in \mathcal{S}$. then from $\infty \notin \{x_n, \ldots, x_{m+1}\}$ we have $\infty \notin \{x_n, \ldots, x_m\} \wedge x_{m+1} \neq \infty$ and then $\sum_{i=n}^{m+1} x_i = \sum_{i=n}^{m} x_i + x_{m+1} \underset{m \in \mathcal{S} \Rightarrow \sum_{i=n}^{m} x_i < \infty \wedge x_{m+1} < \infty}{=} \sum_{i=n}^{m} x_i + x_{m+1} < \infty$ proving that $m + 1 \in \mathcal{S}$

   so by mathematical induction we have that $\mathcal{S} = \{n, \ldots, \infty\}$ proving our theorem. □

We extend now the sum of constant values (see 10.31 for the real case).

Proposition 17.110. *Let* $\{x_i\}_{i \in \{k, \ldots, n\}}$ *be a finite family with* $\forall i \in \{k, \ldots, n\}$ $x_i = c \in \bar{\mathbb{R}}_+$ *then* $\sum_{i=k}^{n} x_i = (n - k + 1) \cdot c$*. More general if* $\{x_i\}_{i \in I}$ *is a family with finite support such that* $\forall i \in I$ *we have* $x_i = c$ *then* $\sum_{i \in I} x_i = \#(I) \cdot c$

Proof. We prove the first part by induction so let $S = \{n \in \{k, \ldots, \infty\}|$if $\{x_i\}_{i \in \{k, \ldots, n\}}$ is a family with $x_i = c$ then $\sum_{i=k}^{n} x_i = n \cdot c\}$ then we have

$k \in S$. If $\{x_i\}_{i \in \{k, \ldots, k\}}$ is such that $x_i = c$ then $\sum_{k=1}^{k} x_i = x_k = c = (k - k + 1) \cdot c$ proving that $k \in S$

$n \in S \Rightarrow n + 1 \in S$. take $\{x_i\}_{i \in \{k, \ldots, n+1\}}$ with $x_i = c$ then we have $\sum_{i=k}^{n+1} x_i = (\sum_{i=k}^{n} x_i) + x_{n+1} \underset{n \in S}{=} (n - k + 1) \cdot c + x_{n+1} = (n - k + 1) \cdot c + c$, we have now two cases either $c = \infty$ and then $((n + 1) - k + 1) \cdot \infty = (n - k + 1) \cdot c + c$ or $c < \infty$ and then $(n - k + 1) \cdot c + c = ((n + 1) - k + 1) \cdot c$, this proves $\sum_{i=1}^{n+1} x_i = ((n + 1) - k + 1) \cdot c$ giving $n + 1 \in S$

Mathematical induction completes then the proof.

For the second part either we have $c=0$ and then $\text{support}(\{x_i\}_{i\in I})=\emptyset$ and thus by definition $\sum_{i\in I} x_i = 0 = \#(I)\cdot 0 = \#(I)\cdot c$ or $c>0$ and then $\text{support}(\{x_i\}_{i\in I})=I$, so there exists a bijection $b:\{1,\ldots,\#(I)\}\rightarrow I$ and $\sum_{i\in I} x_i = \sum_{i=1}^{\#(I)} x_{b(i)} = (\#(I)-1+1)\cdot c = \#(I)\cdot c$. □

We now define the concept of infinite sums of the extended positive reals.

Definition 17.111. *Let $\{x_i\}_{i\in\{n,\ldots,\infty\}}$ be a sequence of extended non negative reals then we say that $\sum_{i=n}^{\infty} x_i = x$ if $\{\sum_{i=n}^{m} x_i\}_{m\in\{n,\ldots,\infty\}}$ is convergent and $\lim_{m\rightarrow\infty}(\sum_{i=n}^{m} x_i) = x$ (using the definition of the limit in $\bar{\mathbb{R}}$).*

Theorem 17.112. *Let $\{x_i\}_{i\in\{n,\ldots,\infty\}}$ be a a sequence of extended non negative reals such that $\sum_{i=n}^{\infty} x_i$ is convergent to x, $\alpha\in\mathbb{R}$ then $\sum_{i=n}^{\infty}(\alpha\cdot x)$ is convergent to $\alpha\cdot x$ or $\sum_{i=n}^{\infty}(\alpha\cdot x_i) = \alpha\cdot x$*

Proof. As $\sum_{i=n}^{\infty}(\alpha\cdot x_i) = \lim_{k\rightarrow\infty}(\sum_{i=n}^{k}\alpha_i\cdot x_i) \underset{17.101}{=} \lim_{k\rightarrow\infty}(\alpha\cdot\sum_{i=n}^{k} x_i) \underset{17.89}{=} \alpha\cdot\lim_{k\rightarrow\infty}(\sum_{i=n}^{k} x_i) = \alpha\cdot\sum_{i=k}^{\infty} x_k$. □

As increasing sequences have always a limit the following theorem will ensure that series of positive extended reals always converges.

Theorem 17.113. *If $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ of non negative extended reals [$\forall i\in\{k,\ldots,\infty\}$ we have $0\leqslant x_i\in\bar{\mathbb{R}}$] then $\{\sum_{i=k}^{n} x_i\}_{n\in\{k,\ldots,\infty\}}$ is a increasing sequence of extended reals. So there exists a $s\in\bar{\mathbb{R}}$ such that $\sum_{i=k}^{\infty} x_i = \lim_{n\rightarrow\infty}(\sum_{i=k}^{n} x_i) = s$ or $\sum_{i=k}^{n} x_i\uparrow s$ (see 17.84) and using 17.83 we have $\sum_{i=k}^{\infty} x_i = \sup(\sum_{i=k}^{n} x_i|n\in\mathbb{N})$*

Proof. This follows from 17.100 and 17.83 □

Theorem 17.114. *Let $\{x_i\}_{i\in\{k,\ldots,\infty\}}$, $\{y_i\}_{i\in\{k,\ldots,\infty\}}$ be sequences of non negative extended reals then $\sum_{i=k}^{\infty}(x_i+y_i) = \sum_{i=k}^{\infty} x_i + \sum_{i=k}^{\infty} y_i$*

Proof. Let $n\in\{k,\ldots,\infty\}$ then we have $\sum_{i=k}^{n}(x_i+y_i) \underset{10.53}{=} \sum_{i=k}^{n} x_i + \sum_{i=k}^{n} y_i$. Next

$$\begin{aligned}\sum_{i=k}^{\infty}(x_i+y_i) &= \lim_{n\rightarrow\infty}\sum_{i=k}^{n}(x_i+y_i)\\ &\underset{17.89}{=} \lim_{n\rightarrow\infty}\sum_{i=k}^{n} x_i + \lim_{n\rightarrow\infty}\sum_{i=k}^{n} y_i\\ &= \sum_{i=k}^{\infty} x_i + \sum_{i=k}^{\infty} y_i\end{aligned}$$

□

Series preserve the order relation in the set of extended reals as the following theorem shows.

Theorem 17.115. *Let $\{x_i\}_{i\in\{k,\ldots,\infty\}}$, $\{y_i\}_{i\in\{k,\ldots,\infty\}}$ be sequences of positive extended reals such that $\forall i\in\{k,\ldots,\infty\}$ we have $x_i\leqslant y_i$ then we have that $\sum_{i=k}^{\infty} x_i\leqslant \sum_{i=k}^{\infty} y_i$*

Proof. Note that $\forall l\in\{k,\ldots,\infty\}$ we have by 10.33 that $\sum_{i=k}^{l} x_i\leqslant\sum_{i=k}^{l} y_i$ and thus by 2.172 we have that $\sum_{i=k}^{\infty} x_i=\sup\left(\left\{\sum_{i=k}^{l} x_i|l\in\{k,\ldots,\infty\}\right\}\right)\leqslant\sup\left(\left\{\sum_{i=k}^{l} y_i|l\in\{k,\ldots,\infty\}\right\}\right)=\sum_{i=k}^{\infty} y_i$ □

The following two theorems are intuitive very clear but still needs some proof.

Theorem 17.116. *Let $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ be a sequence of positive extended reals such that $\sum_{i=k}^{\infty} x_i<\infty$ then $\forall i\in\{k,\ldots,\infty\}$ we have $x_i<\infty$.*

Proof. As $\sup\left(\sum_{i=k}^{n} x_i|n\in\{k,\ldots,\infty\}\right)=\sum_{i=k}^{\infty} x_i<\infty$ we have, using 17.11, that $\forall n\in\{k,\ldots,\infty\}\vDash\sum_{i=k}^{n} x_i<\infty \underset{17.98}{\Rightarrow} x_n<\infty$ □

Theorem 17.117. *Let $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ a sequence of positive extended reals such that $\exists m\in\{k,\ldots,\infty\}$ with $\forall i\in\{m+1,\ldots,\infty\}$ we have $x_i=0$ then $\sum_{i=k}^{\infty} x_i=\sum_{i=k}^{m} x_i$*

Proof. $\sum_{i=k}^{\infty} x_i = \sup\left(\left\{\sum_{i=k}^{n} x_i|n\in\{k,\ldots,\infty\}\right\}\right) \underset{\text{if } n>m\Rightarrow\sum_{i=k}^{n}x_i=\sum_{i=k}^{m}x_i \text{ the biggest sum}}{=} \sum_{i=k}^{m} x_i$ □

We show now that the finite sum of a family of extended positive reals is independent of the order of the terms.

Theorem 17.118. *If $\{x_i\}_{i\in\{n,\ldots,m\}}$ is a finite family of extended reals such that $\{-\infty,\infty\}\nsubseteq\{x_i|i\in\{n,\ldots,\infty\}\}$ and $\sigma:\{n,\ldots,m\}\rightarrow\{n,\ldots,m\}$ then $\sum_{i=n}^{m} x_i=\sum_{i=n}^{m} x_{\sigma(i)}$*

Proof. This follows from the fact that $\langle\bar{\mathbb{R}}_+,+\rangle$ is a Abelian semi-group and $\sum_{i=n}^{m} x_i \underset{17.98}{=} \sum_{i=0}^{m-n} x_{i+n} \underset{10.19}{=} \sum_{i=0}^{m-n} x_{\sigma(i+n)}=\sum_{i=n}^{m} x_{\sigma(i)}$ □

As for infinite sums, the infinite sum of extended positive reals exists always and is independent of the order as is shown in the following theorem.

Theorem 17.119. *If $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ is a sequence of positive extended reals [$\forall i\in\{k,\ldots,\infty\}$ we have $x_i\geqslant 0$] then $\sum_{i=k}^{\infty} x_i$ and $\sum_{i=k}^{\infty} x_{\sigma(i)}$ converges and $\sum_{i=k}^{\infty} x_i=\sum_{i=k}^{\infty} x_{\sigma(i)}$ for every permutation $\sigma:\{k,\ldots,\infty\}\rightarrow\{k,\ldots,\infty\}$*

Proof. As by 17.2 we have that $\{\sum_{i=k}^{n} x_i\}_{n\in\{1,\ldots,\infty\}}$ and $\{\sum_{i=k}^{n} x_i\}_{n\in\{1,\ldots,\infty\}}$ are increasing sequences we have by 17.83 that $\sum_{i=1}^{\infty} x_i$, $\sum_{i=1}^{\infty} x_{\sigma(i)}$ always converges and $\sum_{i=k}^{\infty} x_i=\sup\left(\left\{\sum_{i=1}^{n} x_i|n\in\{k,\ldots,\infty\}\right\}\right)$, $\sum_{i=k}^{\infty} x_{\sigma(i)}=\sup\left(\left\{\sum_{i=1}^{n} x_{\sigma(i)}|n\in\{k,\ldots,\infty\}\right\}\right)$. Next we prove that

$$\text{If } \sigma:\{k,\ldots,\infty\}\rightarrow\{k,\ldots,\infty\} \text{ is a permutation then } \sum_{i=k}^{\infty} x_{\sigma(i)}\leqslant\sum_{i=k}^{\infty} x_i \tag{17.30}$$

Let $s \in \{\sum_{i=k}^{n} x_{\sigma(i)} | n \in \{k, \ldots, \infty\}\}$ then $s = \sum_{i=k}^{n} x_{\sigma(i)}$, define now $m = \max(\{\sigma(i) | i \in \{k, \ldots, n\}\})$ then $\forall i \in \{k, \ldots, n\}$ we have $1 \leqslant \sigma(i) \leqslant m \Rightarrow \sigma(\{k, \ldots, n\}) \subseteq \{k, \ldots, m\}$ then we have $\sum_{i=k}^{m} x_i \underset{10.41}{=} \sum_{i \in \{k, \ldots, m\}} x_i \geqslant_{17.100} \sum_{i \in \sigma(\{k, \ldots, n\})} x_i = \sum_{i \in \{k, \ldots, n\}} x_{\sigma(i)}$, so using 2.172 and the fact that $\sum_{i=k}^{m} x_i \in \{\sum_{k=1}^{n} x_i | n \in \{k, \ldots, \infty\}\}$, we have $\sup(\{\sum_{i=k}^{n} x_{\sigma(i)} | n \in \{1, \ldots, \infty\}\}) \leqslant \sup(\{\sum_{i=k}^{n} x_i | i \in \{k, \ldots, n\}\})$ and thus $\sum_{i=k}^{\infty} x_{\sigma(i)} \leqslant \sum_{i=k}^{\infty} x_i$ as must be proved.

Now if $\sigma: \{k, \ldots, \infty\} \rightarrow \{k, \ldots, \infty\}$ is a bijection take then σ^{-1} and then we have by 17.30 that $\sum_{i=k}^{\infty} x_i = \sum_{i=k}^{\infty} x_{\sigma^{-1}(\sigma(i))} \leqslant \sum_{i=k}^{\infty} x_{\sigma(i)}$ which by 17.30 proves that $\sum_{i=1}^{\infty} x_i = \sum_{i=1}^{\infty} x_{\sigma(i)}$ □

The next theorem shows how we can remove zeroes out of a infinite sum of extended reals.

Theorem 17.120. *Let $k \in \mathbb{N}$ and $\{x_i\}_{i \in \{k, \ldots, \infty\}}$ be a family of extended positive reals and $S = \{i \in \{k, \ldots, n\} | x_i \neq 0\}$ then we have the following possibilities*

- ***$S = \emptyset$.*** *then $\sum_{i=k}^{\infty} x_i = 0$*
- ***$\emptyset \neq S$ is finite.*** *then there exists a bijection $\beta: \{1, \ldots, n\} \rightarrow S$ such that $\sum_{i=k}^{\infty} x_i = \sum_{i=1}^{n} x_{\beta(i)}$*
- ***S is infinite countable.*** *then there exists a bijection $\beta: \mathbb{N} \rightarrow S$ such that $\sum_{i=k}^{\infty} x_i = \sum_{i=1}^{\infty} x_{\beta(i)}$*

Proof. As $S \subseteq \{k, \ldots, \infty\}$ we have that

$$\forall i \in \{k, \ldots, \infty\} \setminus S \vDash x_i = 0 \tag{17.31}$$

For S we have the following cases to consider

- **$S = \emptyset$.** Then $\forall i \in \{k, \ldots, \infty\}$ we have $x_i = 0$. so that $\sum_{i=k}^{\infty} x_i = \sup(\{\sum_{i=k}^{n} x_i | n \in \{k, \ldots, \infty\}\}) = \sup(\{0\}) = 0$
- **$\emptyset \neq S$ is finite.** Then there exists a $n \in \mathbb{N}$ and a bijection $\beta: \{1, \ldots, n\} \rightarrow S$. Take $m = \max(S) \subseteq \{k, \ldots, \infty\}$ then if $i > m$ we have $i \notin S$ and thus by 17.31 $x_i = 0$ giving

$$\forall i > m \text{ we have } x_i = 0 \tag{17.32}$$

Further $S \subseteq \{k, \ldots, m\}$ so that

$$\{k, \ldots, m\} = (\{k, \ldots, m\} \setminus S) \bigcup S \text{ and } \{k, \ldots, m\} \setminus S \bigcap S = \emptyset \tag{17.33}$$

Now for $l \in \{k, \ldots, \infty\}$ we have then the following cases to consider in relation to m

$l \leqslant m$. then

$$\begin{aligned}
\sum_{i=k}^{l} x_i &= \sum_{i \in \{k, \ldots, m\}} x_i \\
&\underset{10.45 \text{ and } 17.33}{=} \sum_{i \in \{k, \ldots, m\} \setminus S} x_i + \sum_{i \in S} x_i \\
&\underset{17.31}{=} 0 + \sum_{i \in S} x_i
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i\in\beta(\{1,\ldots,n\})} x_i \\
&\underset{10.44}{=} \sum_{i\in\{1,\ldots,n\}} x_{\beta(i)} \\
&= \sum_{i=1}^{n} x_{\beta(i)}
\end{aligned}$$

$l > m$. then

$$\begin{aligned}
\sum_{i=k}^{l} x_i &= \sum_{i=k}^{m} x_i + \sum_{i=m+1}^{l} x_i \\
&\underset{17.31}{=} \sum_{i=k}^{m} x_i \\
&= \sum_{i\in\{k,\ldots,m\}} x_i \\
&= \sum_{i\in\beta(\{1,\ldots,n\})} x_i \\
&\underset{10.44}{=} \sum_{i\in\{1,\ldots,n\}} x_{\beta(i)} \\
&= \sum_{i=1}^{n} x_{\beta(i)}
\end{aligned}$$

So it follows that $\forall l \in \{k,\ldots,\infty\}$ that $\sum_{i=k}^{\infty} x_i = \sum_{i=1}^{n} x_{\beta(i)}$. Hence $\sum_{i=k}^{\infty} x_i = \sup(\{\sum_{i=k}^{l} x_i | l \in \{k,\ldots,\infty\}\}) = \sup(\{\sum_{i=k}^{n} x_{\beta(i)}\}) = \sum_{i=k}^{n} x_{\beta(i)}$ proving that

$$\sum_{i=k}^{\infty} x_i = \sum_{i=k}^{n} x_{\beta(i)}$$

S is infinite countable. Then there exists a bijection $\beta: \mathbb{N} \to S$. Take $n \in \{k,\ldots,\infty\}$ then for $\{k,\ldots,n\}$ we can have either

$\forall i \in \{k,\ldots,n\} \vDash x_i = 0$. then $\sum_{i=1}^{n} x_i = 0 \leqslant \sum_{i=1}^{1} x_{\beta(i)}$

$\exists i \in \{k,\ldots,n\} \vDash x_i \neq 0$. then $i \in \{k,\ldots,n\} \bigcap S$ and as β is surjective there exists a $j \in \{1,\ldots,\infty\}$ such that $\beta(j) = i \in \{k,\ldots,n\} \bigcap S$ proving that $\emptyset \neq \beta^{-1}(\{k,\ldots,n\} \bigcap S)$. As $\beta^{-1}(\{k,\ldots,n\} \bigcap S)$ is also finite $m = \max(\beta^{-1}(\{k,\ldots,n\} \bigcap S))$ exists, then $\beta^{-1}(\{k,\ldots,n\} \bigcap S) \subseteq \{1,\ldots,m\}$ so that

$$\{k,\ldots,n\} \bigcap S \subseteq \beta(1,\ldots,m) \tag{17.34}$$

Futher we have that

$$\{k,\ldots,n\} = (\{k,\ldots,n\} \setminus S) \bigcup (\{k,\ldots,n\} \bigcap S) \tag{17.35}$$

and

$$(\{k,\ldots,n\} \setminus S) \bigcap (\{k,\ldots,n\} \bigcap S) \tag{17.36}$$

So

$$\begin{array}{lll}
\sum_{i=k}^{n} x_i & = & \sum_{i\in\{k,\ldots,n\}} x_i \\
& \underset{10.45 \text{ and } 17.35,17.36}{=} & \sum_{i\in\{k,\ldots,n\}\setminus S} x_i + \sum_{i\in\{k,\ldots,n\}\bigcap S} x_i \\
& = & 0 + \sum_{i\in\{k,\ldots,n\}\bigcap S} x_i \\
& \leqslant & \sum_{i\in\{k,\ldots,n\}\bigcap S} x_i + \sum_{\beta\{1,\ldots,m\}\setminus(\{k,\ldots,n\}\bigcap S)} x_i \\
& \underset{10.45 \text{ and } 17.34}{=} & \sum_{i\in\beta(\{1,\ldots,m\})} x_i \\
& \underset{10.44}{=} & \sum_{i\in\{1,\ldots,m\}} x_{\beta(i)} \\
& = & \sum_{i=1}^{m} x_{\beta(i)}
\end{array}$$

Hence $\forall n\in\{k,\ldots,\infty\}$ there exists a $m\in\{1,\ldots,\infty\}$ such that $\sum_{i=1}^{n} x_i \leqslant \sum_{i=1}^{m} x_i$ proving that

$$\sup\left(\left\{\sum_{i=k}^{n} x_i | n\in\{k,\ldots,\infty\}\right\}\right) \leqslant \sup\left(\left\{\sum_{i=1}^{n} x_{\beta(i)} | n\in\{1,\ldots,\infty\}\right\}\right) \tag{17.37}$$

For the opposite inequality, take $n\in\{1,\ldots,\infty\}$ then $\beta(\{1,\ldots,n\})$ is non empty and finite so that $m=\max(\beta(\{1,\ldots,n\}))\subseteq S\subseteq\{k,\ldots,\infty\}$ exists giving

$$\beta(\{1,\ldots,n\})\subseteq\{k,\ldots,m\} \tag{17.38}$$

so that

$$\begin{array}{lll}
\sum_{i=1}^{n} x_{\beta(i)} & = & \sum_{i\in\{1,\ldots,n\}} x_{\beta(i)} \\
& \underset{10.44}{=} & \sum_{i\in\beta(\{1,\ldots,n\})} x_i \\
& \leqslant & \sum_{i\in\{k,\ldots,m\}\setminus\beta(\{1,\ldots,n\})} x_i + \sum_{i\in\beta(\{k,\ldots,m\})} x \\
& \underset{10.45 \text{ and } 17.38}{=} & \sum_{i\in\{k,\ldots,m\}} x_i \\
& = & \sum_{i=k}^{m} x_i
\end{array}$$

Hence $\forall n\in\{1,\ldots,\infty\}$ there exists a $m\in\{k,\ldots,\infty\}$ such that $\sum_{i=1}^{n} x_i \leqslant \sum_{i=k}^{m} x_i$ proving

$$\sup\left(\left\{\sum_{i=1}^{n} x_{\beta(i)} | n\in\{1,\ldots,\infty\}\right\}\right)\leqslant \sup\left(\left\{\sum_{i=k}^{n} x_i | n\in\{k,\ldots,\infty\}\right\}\right)$$

Finally using the above together with 17.37 we have

$$\sum_{i=k}^{\infty} x_i = \sup\left(\left\{\sum_{i=k}^{n} x_i | n\in\{k,\ldots,\infty\}\right\}\right) = \sup\left(\left\{\sum_{i=1}^{n} x_{\beta(i)} | n\in\{1,\ldots,\infty\}\right\}\right) = \sum_{i=1}^{\infty} x_{\beta(i)}$$

□

The next theorem shows that we can interchange infinite sums.

Theorem 17.121. *Let $k,l\in\mathbb{N}$, $\{x_{(i,j)}\}_{(i,j)\in\{k,\ldots,\infty\}\times\{l,\ldots,\infty\}}$ be a family of extended positive reals then $\sum_{i=k}^{\infty}(\sum_{j=l}^{\infty} x_{(i,j)}) = \sum_{j=l}^{\infty}(\sum_{i=k}^{\infty} x_{(i,j)})$. Note that using 17.117 we have also for $\{x_{(i,j)}\}_{(i,j)\in\{k,\ldots,\infty\}\times\{l,\ldots,n\}}$ that $\sum_{i=k}^{\infty}(\sum_{j=l}^{n} x_{(i,j)}) = \sum_{j=l}^{n}(\sum_{i=k}^{\infty} x_{(i,j)})$ and using 17.117 we have for $\{x_{(i,j)}\}_{(i,j)\in\{k,\ldots,n\}\times\{l,\ldots,m\}}$ that $\sum_{i=k}^{n}(\sum_{j=l}^{m} x_{(i,j)}) = \sum_{j=l}^{m}(\sum_{i=k}^{n} x_{(i,j)})$.*

Proof. Let $n\in\{k,\ldots,\infty\}$, $m\in\{l,\ldots,\infty\}$ then $\sum_{i=k}^{n}(\sum_{j=l}^{m} x_{(i,j)}) \underset{10.41}{=} \sum_{i\in\{k,\ldots,n\}}(\sum_{j\in\{l,\ldots,m\}} x_{(i,j)}) \underset{10.48}{=} \sum_{j\in\{l,\ldots,m\}}(\sum_{i\in\{k,\ldots,n\}} x_{(i,j)}) = \sum_{j=l}^{m}(\sum_{i=k}^{n} x_{(i,j)}) \leqslant \sum_{j=l}^{m}(\sup(\{\sum_{i=k}^{n} x_{(i,j)} | n\in\{l,\ldots,\infty\}\})) = \sum_{j=l}^{m}(\sum_{i=k}^{\infty} x_{(i,j)}) \leqslant \sup(\{\sum_{j=l}^{n}(\sum_{i=k}^{\infty} x_{(i,j)}) | n\in\{l,\ldots,m\}\}) = \sum_{j=l}^{\infty}(\sum_{i=k}^{\infty} x_{(i,j)})$ from which follows $\sup(\{\sum_{i=k}^{n}(\sum_{j=l}^{m} x_{(i,j)} | n\in\{k,\ldots,\infty\})\}) \leqslant \sum_{j=l}^{\infty}(\sum_{i=k}^{\infty} x_{(i,j)}) \underset{2.172}{\Rightarrow} \sup(\{\sum_{i=k}^{n}(\sup(\{\sum_{j=l}^{m} x_{(i,j)} | m\in\{l,\ldots,\infty\}\})) | n\in\{k,\ldots,\infty\}\}) \leqslant \sum_{j=l}^{\infty}(\sum_{i=k}^{\infty} x_{(i,j)})$ giving

$$\sum_{i=k}^{\infty}\left(\sum_{j=l}^{\infty} x_{(i,j)}\right) \leqslant \sum_{j=l}^{\infty}\left(\sum_{i=k}^{\infty} x_{(i,j)}\right) \tag{17.39}$$

Likewise if $n\in\{k,\ldots,\infty\}$, $m\in\{l,\ldots,\infty\}$ then $\sum_{j=l}^{n}(\sum_{i=k}^{m} x_{(i,j)}) \underset{10.41}{=} \sum_{j\in\{l,\ldots,n\}}(\sum_{i\in\{k,\ldots,m\}} x_{(i,j)}) \underset{10.48}{=} \sum_{i\in\{k,\ldots,m\}}(\sum_{j\in\{l,\ldots,n\}} x_{(i,j)}) = \sum_{i=k}^{m}(\sum_{j=l}^{n} x_{(i,j)}) \leqslant \sum_{i=k}^{m}(\sup(\{\sum_{j=l}^{n} x_{(i,j)} | n\in\{l,\ldots,\infty\}\})) = \sum_{i=k}^{m}(\sum_{j=l}^{\infty} x_{(i,j)}) \leqslant \sup(\{\sum_{i=k}^{n}(\sum_{j=l}^{\infty} x_{(i,j)}) | n\in\{l,\ldots,m\}\}) = \sum_{i=k}^{\infty}(\sum_{j=l}^{\infty} x_{(i,j)})$ from which follows $\sup(\{\sum_{j=l}^{n}(\sum_{i=k}^{m} x_{(i,j)} | n\in\{l,\ldots,\infty\})\}) \leqslant$

$\sum_{i=k}^{\infty}\left(\sum_{j=l}^{\infty} x_{(i,j)}\right) \underset{2.172}{\Rightarrow} \sup\left(\left\{\sum_{j=l}^{n}\left(\sup\left(\left\{\sum_{i=k}^{m} x_{(i,j)} | m \in \{k, \ldots, \infty\}\right\}\right)\right) | n \in \{l, \ldots, \infty\}\right\}\right) \leqslant \sum_{i=k}^{\infty}\left(\sum_{j=l}^{\infty} x_{(i,j)}\right)$ giving

$$\sum_{j=l}^{\infty}\left(\sum_{i=k}^{\infty} x_{(i,j)}\right) \leqslant \sum_{i=k}^{\infty}\left(\sum_{j=l}^{\infty} x_{(i,j)}\right) \tag{17.40}$$

17.39 and 17.40 proves then finally the theorem. □

Every finite sum of a denumerable sums can be writen as a denumerable sum

Theorem 17.122. *Let $n \in \{l, \ldots, \infty\}$ and $\{x_{(i,j)}\}_{(i,j) \in \{k, \ldots, n\} \times [l, \ldots, \infty]}$ be a family of non negative extended real then if $\sigma: \{p, \ldots, \infty\} \rightarrow \{k, \ldots, n\} \times \{l, \ldots, \infty\}$ is a bijection (which exist by 5.64) then $\sum_{j=l}^{\infty}\left(\sum_{i=k}^{n} x_{(i,j)}\right) \underset{17.121}{=} \sum_{i=k}^{n}\left(\sum_{j=l}^{\infty} x_{(i,j)}\right) = \sum_{i=p}^{\infty} x_{\sigma(i)} = \sum_{i=p}^{\infty} x_{(\sigma(i)_1, \sigma(i)_2)}$*

Proof. Let $m \in \{p, \ldots, \infty\}$ and define $N_m = \max(\{\sigma(i)_2 | i \in \{p, \ldots, m\}\})$ then we have $\sigma(\{p, \ldots, m\}) \subseteq \{k, \ldots, n\} \times \{l, \ldots, N_m\}$ so that

$$\begin{aligned}
\sum_{i=p}^{m} x_{\sigma(i)} & \underset{10.41}{=} \sum_{i \in \{p, \ldots, m\}} x_{\sigma(i)} \\
& \underset{10.44}{=} \sum_{(i,j) \in \sigma(\{p, \ldots, m\})} x_{(i,j)} \\
& \leqslant_{17.100} \sum_{(i,j) \in \{k, \ldots, n\} \times \{l, \ldots, N_m\}} x_{(i,j)} \\
& \underset{10.48}{=} \sum_{i \in \{k, \ldots, n\}} \left(\sum_{j \in \{l, \ldots, N_m\}} x_{(i,j)}\right) \\
& \underset{10.41}{=} \sum_{i=k}^{n}\left(\sum_{j=l}^{N_m} x_{(i,j)}\right) \\
& \leqslant_{10.33} \sum_{i=k}^{n} \sup\left(\left\{\sum_{j=l}^{m} x_{(i,j)} | m \in \mathbb{N}\right\}\right) \\
& = \sum_{i=k}^{n}\left(\sum_{j=l}^{\infty} x_{(i,j)}\right)
\end{aligned}$$

From the above it follows that $\sum_{i=p}^{\infty} x_{\sigma(i)} = \sup\left(\left\{\sum_{i=p}^{m} x_{\sigma(i)} | m \in \mathbb{N}\right\}\right) \leqslant \sum_{i=k}^{n}\left(\sum_{j=l}^{\infty} x_{(i,j)}\right)$ so we have

$$\sum_{i=p}^{\infty} x_{\sigma(i)} \leqslant \sum_{i=k}^{n}\left(\sum_{j=l}^{\infty} x_{(i,j)}\right) \tag{17.41}$$

For the opposite inequality let $m \in \{l, ..., \infty\}$ and take $N_m = \max(\sigma^{-1}(\{k, ..., n\} \times \{l, ..., m\}))$ then if $i \in \sigma^{-1}(\{k, ..., n\} \times \{l, ..., m\}) \Rightarrow p \leqslant i \leqslant N_m$ proving that $\sigma^{-1}(\{k, ..., n\} \times \{l, ..., m\}) \subseteq \{p, ..., N_m\}$. So $\{k, ..., n\} \times \{l, ..., m\} \underset{\sigma \text{ is a bijection}}{=} \sigma(\sigma^{-1}(\{k, ..., n\} \times \{l, ..., m\})) \subseteq \sigma(\{p, ..., N_m\})$. It follows that

$$\begin{aligned}
\sum_{i=k}^{n}\left(\sum_{j=l}^{m} x_{(i,j)}\right) &\underset{10.48}{=} \sum_{(i,j)\in\{k,...,n\}\times\{l,...,m\}} x_{(i,j)} \\
&\leqslant_{17.100} \sum_{(i,j)\in\sigma(\{p,...,N_m\})} x_{(i,j)} \\
&\underset{10.44}{=} \sum_{i\in\{p,...,N_m\}} x_{\sigma(i)} \\
&= \sum_{i=p}^{N_m} x_{\sigma(i)} \\
&\leqslant \sup\left(\left\{\sum_{i=p}^{m} x_{\sigma(i)} | m \in \mathbb{N}\right\}\right) \\
&= \sum_{i=p}^{\infty} x_{\sigma(i)}
\end{aligned}$$

Then we have

$$\begin{aligned}
\sum_{i=k}^{n}\left(\sum_{j=l}^{\infty} x_{(i,j)}\right) &\underset{17.121}{=} \sum_{j=l}^{\infty}\left(\sum_{i=k}^{n} x_{(i,j)}\right) \\
&= \sup\left\{\left(\sum_{j=l}^{m}\left(\sum_{i=k}^{n} x_{(i,j)}\right)\right) | m \in \mathbb{N}\right\} \\
&\underset{17.121}{=} \sup\left\{\left(\sum_{i=k}^{n}\left(\sum_{j=l}^{m} x_{(i,j)}\right)\right) | m \in \mathbb{N}\right\} \\
&\leqslant \sum_{i=p}^{\infty} x_{\sigma(i)}
\end{aligned}$$

proving together with 17.41 that $\sum_{i=1}^{\infty} x_{\sigma(i)} = \sum_{i=k}^{n} \left(\sum_{j=l}^{\infty} x_{(i,j)}\right)$ and thus the theorem. $\square$

Every denumerable sum of finite sums can be written as a denumerable sun.

Theorem 17.123. *Let* $\{x_{(i,j)}\}_{(i,j)\in\bigsqcup_{i\in\{k,...,\infty\}}\{(i,m_i),...,(i,n_i)\}}$ *[where* $\forall i \in \{k, ..., \infty\}$ *we have* $m_i \leqslant n_i$*] we have then if* $\sigma: \{l, ..., \infty\} \rightarrow \bigsqcup_{i\in\{k,...,\infty\}} \{(i, m_i), ..., (i, n_i)\}$ *is a bijection (which must exists by 5.72) we have* $\sum_{i=l}^{\infty} x_{\sigma(i)} = \sum_{i=k}^{\infty} \left(\sum_{i=m_i}^{n_i} x_{(i,j)}\right)$

Proof. First let $i \in \{k, \ldots, \infty\}$ and define the bijection $\tau_i: \{m_i, \ldots, n_i\} \rightarrow \{(i, m_i), \ldots, (i, n_i)\}$ by $\tau_i(k) = (i, k)$ so that by 10.44 we have $\sum_{(i,j) \in \{(i,1), \ldots, (i,n_i)\}} x_{(i,j)} = \sum_{k \in \{m_i, \ldots, n_i\}} x_{\tau_i(k)} = \sum_{k \in \{m_i, \ldots, n_i\}} x_{(i,k)} = \sum_{k=m_i}^{n_i} x_{(i,k)}$ proving that

$$\sum_{(i,j) \in \{(i,m_i), \ldots, (i,n_i)\}} x_{(i,j)} = \sum_{j \in \{m_i, \ldots, n_i\}} x_{(i,j)} = \sum_{j=m_i}^{n_j} x_{(i,j)} \tag{17.42}$$

Let $m \in \{l, \ldots, \infty\}$ define then $N_m = \max(\{\sigma(i)_1 | i \in \{l, \ldots, m\}\})$. If $i \in \{l, \ldots, m\}$ then $\sigma(i) \in \bigsqcup_{i \in \{k, \ldots, \infty\}} \{(i, m_i), \ldots, (i, n_i)\}$ there exists a $j \in \{k, \ldots, \infty\}$ such that $\sigma(i) \in \{(j, m_j), \ldots, (j, n_j)\}$ so that as $\sigma(i)_1 = j \in \{k, \ldots, N_m\}$ proving that

$$\sigma(\{l, \ldots, m\}) \subseteq \bigsqcup_{i \in \{k, \ldots, N_m\}} \{(i, m_i), \ldots, \ldots, (i, n_i)\} \text{ (a pairwise disjoint union)} \tag{17.43}$$

Next

$$\begin{aligned}
\sum_{i=l}^{m} x_{\sigma(i)} &\underset{10.41}{=} \sum_{i \in \{l, \ldots, m\}} x_{\sigma(i)} \\
&= \sum_{(i,j) \in \sigma(\{l, \ldots, m\})} x_{(i,j)} \\
&\leqslant_{17.43 \text{ and } 17.100} \sum_{(i,j) \in \bigcup_{r \in \{k, \ldots, N_m\}} \{(r, m_r), \ldots, (r, n_r)\}} x_{(i,j)} \\
&\underset{10.46}{=} \sum_{r \in \{k, \ldots, N_m\}} \left(\sum_{(i,j) \in \{(r, m_r), \ldots, (r, n_r)\}} x_{(i,j)} \right) \\
&= \sum_{r=k}^{N_m} \left(\sum_{(i,j) \in \{(r, m_r), \ldots, (r, n_r)\}} x_{(i,j)} \right)
\end{aligned}$$

proving using the above and 17.42 that

$$\sum_{i=l}^{m} x_{\sigma(i)} \leqslant \sum_{r=k}^{N_m} \left(\sum_{j=m_r}^{n_r} x_{(r,j)} \right) \tag{17.44}$$

Now $\forall m \in \mathbb{N}$ we have $\sum_{r=k}^{N_m} (\sum_{j=m_r}^{n_r} x_{(r,j)}) \leqslant \sup(\{\sum_{r=k}^{n} (\sum_{j=m_r}^{n_r} x_{(r,j)}) | n \in \mathbb{N}\}) = \sum_{r=1}^{\infty} (\sum_{j=m_r}^{n_r} x_{(r,j)})$ proving by the definition of a supremum together with 17.44 that

$$\sum_{i=l}^{\infty} x_{\sigma(i)} \leqslant \sum_{r=k}^{\infty} \left(\sum_{j=m_r}^{n_r} x_{(r,j)} \right) \tag{17.45}$$

For the opposite inequality take $m \in \{k, \ldots, \infty\}$ and take $N_m = \max\left(\sigma^{-1}\left(\bigsqcup_{i \in \{k, \ldots, m\}} \{(i, m_i), \ldots, (i, n_i)\}\right)\right)$ then if $i \in \sigma^{-1}\left(\bigcup_{j \in \{k, \ldots, m\}} \{(j, m_j), \ldots, (j, n_j)\}\right)$ then $i \in \{l, \ldots, N_m\}$ so that $\sigma^{-1}\left(\bigsqcup_{j \in \{k, \ldots, m\}} \{(j, m_j), \ldots, (j, n_j)\}\right) \subseteq \{l, \ldots, N_m\}$ giving as σ is a bijection that

$$\bigsqcup_{j \in \{k, \ldots, m\}} \{(j, m_j), \ldots, (j, n_j)\} \subseteq \sigma(\{l, \ldots, N_m\}) \tag{17.46}$$

Now we have

$$\begin{aligned}
\sum_{i=k}^{m}\left(\sum_{j=m_i}^{n_i} x_{(i,j)}\right) &\underset{17.42}{=} \sum_{r=k}^{m}\left(\sum_{(i,j)\in\{(r,m_r),\ldots,(r,n_r)\}} x_{(i,j)}\right)\\
&\underset{10.46}{=} \sum_{(i,j)\in\bigsqcup_{r\in\{k,\ldots,m\}}\{(r,m_r),\ldots,(r,n_r)\}} x_{(i,j)}\\
&\leqslant_{17.100 \text{ together with } 17.46} \sum_{(i,j)\in\sigma(\{l,\ldots,N_m\})} x_{(i,j)}\\
&\underset{10.44}{=} \sum_{i\in\{l,\ldots,N_m\}} x_{\sigma(i)} = \sum_{i=l}^{N_m} x_{\sigma(i)}\\
&\leqslant \sup\left(\left\{\sum_{i=l}^{n} x_{\sigma(i)} | n\in\mathbb{N}\right\}\right)\\
&= \sum_{i=l}^{\infty} x_{\sigma(i)}
\end{aligned}$$

Using the definition of a supremum we have then $\sum_{i=k}^{\infty}\left(\sum_{j=m_i}^{n_i} x_{(i,j)}\right) = \sup\left(\left\{\sum_{i=1}^{m}\left(\sum_{j=m_i}^{n_i} x_{(i,j)}\right) | m\in\mathbb{N}\right\}\right) \leqslant \sum_{i=l}^{\infty} x_{\sigma(i)}$, this together with 17.45 proves

$$\sum_{i=l}^{\infty} x_{\sigma(i)} = \sum_{i=k}^{\infty}\left(\sum_{j=m_i}^{n_i} x_{(i,j)}\right)$$

□

We prove that a denumerable sum of denumerable sums may be written by as a denumerable sum.

Theorem 17.124. *Let $\{\{x_{i,j}\}_{i\in\mathbb{N}}\}_{j\in\mathbb{N}}$ a sequence of sequences of extended positive reals and $\sigma: \mathbb{N}\rightarrow\mathbb{N}\times\mathbb{N}$ a bijection (which exists as by 5.59 $\mathbb{N}\times\mathbb{N}$ is denumerable) then $\sum_{i\in\mathbb{N}} x_{\sigma(i)_1,\sigma(i)_2} = \sum_{i=1}^{\infty}\left(\sum_{j=1}^{\infty} x_{i,j}\right) \underset{17.121}{=} \sum_{j=1}^{\infty}\left(\sum_{i=1}^{\infty} x_{i,j}\right)$*

Proof. Let $n\in\mathbb{N}$ take then $N_n = \max\left(\{\sigma(i)_1 | i\in\{1,\ldots,n\}\}\right)$, $M_n = \max\left(\{\sigma(i)_2 | i\in\{1,\ldots,n\}\}\right)$ then we have that $\sigma(\{1,\ldots,n\})\subseteq\{1,\ldots,N_n\}\times\{1,\ldots,M_n\}$ and thus

$$\begin{aligned}
\sum_{i=1}^{n} x_{\sigma(i)_1,\sigma(i)_2} &\underset{10.41}{=} \sum_{i\in\{1,\ldots,n\}} x_{\sigma(i)_1,\sigma(i)_2}\\
&\underset{10.44}{=} \sum_{(i,j)\in\sigma\{1,\ldots,n\}} x_{i,j}\\
&\leqslant_{17.100} \sum_{(i,j)\in\{1,\ldots,N_n\}\times\{1,\ldots,M_n\}} x_{i,j}\\
&\underset{10.48}{=} \sum_{i\in\{1,\ldots,N_n\}}\left(\sum_{j\in\{1,\ldots,M_n\}} x_{i,j}\right)\\
&= \sum_{i=1}^{N_n}\left(\sum_{j=1}^{M_n} x_{i,j}\right)
\end{aligned}$$

$$
\begin{aligned}
&\leqslant \sum_{i=1}^{N_n} \sup\left(\left\{\sum_{j=1}^{m} x_{i,j} | m \in \mathbb{N}\right\}\right)\\
&\underset{17.113}{=} \sum_{i=1}^{N_n} \left(\sum_{j=1}^{\infty} x_{i,j}\right)\\
&\leqslant \sup\left(\left\{\sum_{i=}^{n}\left(\sum_{j=1}^{\infty} x_{i,j}\right) | n \in \mathbb{N}\right\}\right)\\
&\underset{17.113}{=} \sum_{i=1}^{\infty}\left(\sum_{j=1}^{\infty} x_{i,j}\right)
\end{aligned}
$$

So $\sup\left(\left\{\sum_{i=1}^{n} x_{\sigma(i)_1,\sigma(i)_2} | n \in \mathbb{N}\right\}\right) \leqslant \sum_{i=1}^{\infty}\left(\sum_{j=1}^{\infty} x_{i,j}\right)$ or using 17.113

$$
\sum_{i \in \{1,\ldots,n\}} x_{\sigma(i)_1,\sigma(i)_2} \leqslant \sum_{i=1}^{\infty}\left(\sum_{j=1}^{\infty} x_{i,j}\right) \tag{17.47}
$$

Let now take $n, m \in \mathbb{N}$ and take $N = \max\left(\sigma^{-1}(\{1,\ldots,n\} \times \{1,\ldots,m\})\right)$ then if $i \in \sigma^{-1}(\{1,\ldots,n\} \times \{1,\ldots,m\})$ we have $1 \leqslant i \leqslant N$ proving that $\sigma^{-1}(\{1,\ldots,n\} \times \{1,\ldots, m\}) \subseteq \{1,\ldots,N\}$. Hence as σ is a bijection we have that $\{1,\ldots,n\} \times \{1,\ldots,m\} \subseteq \sigma(\{1,\ldots,N\})$. So

$$
\begin{aligned}
\sum_{i=1}^{n}\left(\sum_{j=1}^{m} x_{i,j}\right) &\underset{10.41}{=} \sum_{i \in \{1,\ldots,n\}}\left(\sum_{j \in \{1,\ldots,m\}} x_{i,j}\right)\\
&\underset{10.48}{=} \sum_{(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\}} x_{i,j}\\
&\leqslant_{17.100} \sum_{(i,j) \in \sigma(\{1,\ldots,N\})} x_{i,j}\\
&\underset{10.44}{=} \sum_{i=1}^{N} x_{\sigma(i)_1,\sigma(i)_2}\\
&\leqslant \sup\left(\left\{\sum_{i=1}^{k} x_{\sigma(i)_1,\sigma(i)_2} | k \in \mathbb{N}\right\}\right)\\
&\underset{17.113}{=} \sum_{i=1}^{\infty} x_{\sigma(i)_1,\sigma(i)_2}
\end{aligned} \tag{17.48}
$$

Further

$$
\begin{aligned}
\sum_{i=1}^{\infty}\left(\sum_{j=1}^{m} x_{i,j}\right) &\underset{17.113}{=} \sup\left(\left\{\sum_{i=1}^{n}\left(\sum_{j=1}^{m} x_{i,j}\right) | n \in \mathbb{N}\right\}\right)\\
&\leqslant_{17.48} \sum_{i=1}^{\infty} x_{\sigma(i)_1,\sigma(i)_2}
\end{aligned} \tag{17.49}
$$

Switching the sums (see 17.121) we have then that $\sum_{j=1}^{m}\left(\sum_{i=1}^{\infty} x_{i,j}\right) = \sum_{i=1}^{\infty}\left(\sum_{j=1}^{m} x_{i,j}\right) \leqslant_{17.49} \sum_{i=1}^{\infty} x_{\sigma(i)_1,\sigma(i)_2}$ so that $\sum_{j=1}^{\infty}\left(\sum_{i=1}^{\infty} x_{i,j}\right) = \sup\left(\left\{\sum_{j=1}^{n}\left(\sum_{i=1}^{\infty} x_{i,j}\right) | n \in \mathbb{N}\right\}\right) \leqslant \sum_{i=1}^{\infty} x_{\sigma(i)_1,\sigma(i)_2}$. Hence $\sum_{i=1}^{\infty}\left(\sum_{j=1}^{\infty} x_{i,j}\right) \underset{17.121}{=} \sum_{j=1}^{\infty}\left(\sum_{i=1}^{\infty} x_{i,j}\right) \leqslant \sum_{i=1}^{\infty} x_{\sigma(i)_1,\sigma(i)_2}$. Finally this with 17.47 we have

$$\sum_{i=1}^{\infty} x_{\sigma(i)_1,\sigma(i)_2} = \sum_{i=1}^{\infty}\left(\sum_{j=1}^{\infty} x_{i,j}\right) \qquad \square$$

Theorem 17.125. *Let $n \in \mathbb{N}$, $\{\sigma_i: \mathbb{N} \rightarrow \sigma(\mathbb{N})\}_{i \in \{1,\ldots,n\}}$ a set of bijections so that $\bigcup_{i \in \{1,\ldots,n\}} \sigma_i(\mathbb{N}) = \mathbb{N}$ and $\forall i, j \in \{1,\ldots,n\}$ with $i \neq j$ we have $\sigma_i(\mathbb{N}) \bigcap \sigma_j(\mathbb{N}) = \emptyset$ and $\{x_i\}_{i \in \mathbb{N}}$ a sequence of non-negative extended reals then $\sum_{i=1}^{n}\left(\sum_{j=1}^{\infty} x_{\sigma_i(j)}\right) = \sum_{i=1}^{\infty} x_i$*

Proof. First we prove that $\sum_{i=1}^{n}\left(\sum_{j=1}^{\infty} x_{\sigma_i(j)}\right) \leqslant \sum_{i=1}^{\infty} x_i$. Given $m \in \mathbb{N}$ we have $\forall i, j \in \{1,\ldots,n\} \vdash i \neq j$ that $\sigma_i(\{1,\ldots,m\}) \bigcap \sigma_j(\{1,\ldots,m\}) \subseteq \sigma_i(\mathbb{N}) \bigcap \sigma_j(\mathbb{N}) = \emptyset$ and thus applying 10.46 we have

$$\begin{aligned}
\sum_{i=1}^{n}\left(\sum_{j=1}^{m} x_{\sigma_i(j)}\right) &= \sum_{i \in \{1,\ldots,n\}}\left(\sum_{j \in \{1,\ldots,n\}} x_{\sigma_i(j)}\right)\\
&= \sum_{i \in \{1,\ldots,n\}}\left(\sum_{j \in \sigma_i(\{1,\ldots,m\})} x_j\right)\\
&\underset{10.46}{=} \sum_{i \in \bigcup_{j \in \{1,\ldots,n\}} \sigma_j(\{1,\ldots,m\})} x_i
\end{aligned}$$

proving

$$\forall m \in \mathbb{N} \text{ we have } \sum_{i=1}^{n}\left(\sum_{j=1}^{m} x_{\sigma_i(j)}\right) = \sum_{i \in \bigcup_{j \in \{1,\ldots,n\}} \sigma_j(\{1,\ldots,m\})} x_i \tag{17.50}$$

Take now $N_m = \max\left(\bigcup_{j \in \{1,\ldots,n\}} \sigma_j(\{1,\ldots,m\})\right)$ then $\bigcup_{j \in \{1,\ldots,n\}} \sigma_j(\{1,\ldots,m\}) \subseteq \{1,\ldots,N\}$ and by 17.100 we have $\sum_{i \in \bigcup_{j \in \{1,\ldots,n\}} \sigma_j(\{1,\ldots,m\})} x_i \leqslant \sum_{i \in \{1,\ldots,N_m\}} x_i = \sum_{i=1}^{N_m} x_i$. So using 17.50 we have $\sum_{i=1}^{n}\left(\sum_{j=1}^{m} x_{\sigma_i(j)}\right) \leqslant \sum_{i=1}^{N_m} x_i \leqslant \sup\left(\sum_{i=1}^{m} x_i | m \in \mathbb{N}\right) \underset{17.113}{=} \sum_{i=1}^{\infty} x_i$ proving, as by 17.121 we have $\sum_{j=1}^{m}\left(\sum_{i=1}^{n} x_{\sigma_i(j)}\right) = \sum_{i=1}^{n}\left(\sum_{j=1}^{m} x_{\sigma_i(j)}\right)$ that $\sum_{j=1}^{m}\left(\sum_{i=1}^{n} x_{\sigma_i(j)}\right) \leqslant \sum_{i=1}^{\infty} x_i$. Using the definition of the supremum we have then that $\sum_{j=1}^{\infty}\left(\sum_{i=1}^{n} x_{\sigma_i(j)}\right) = \sup\left(\left\{\sum_{j=1}^{m}\left(\sum_{i=1}^{n} x_{\sigma_i(j)}\right) | m \in \mathbb{N}\right\}\right) \leqslant \sum_{i=1}^{\infty} x_i$ giving

$$\sum_{i=1}^{n}\left(\sum_{j=1}^{\infty} x_{\sigma_i(j)}\right) \underset{17.121}{=} \sum_{j=1}^{\infty}\left(\sum_{i=1}^{n} x_{\sigma_i(j)}\right) \leqslant \sum_{i=1}^{\infty} x_i \tag{17.51}$$

For the opposite inequality let $m \in \mathbb{N}$ then $\forall k \in \{1, \ldots, m\}$ we have, as $\mathbb{N} = \bigcup_{i \in \{1, \ldots, n\}} \sigma_i(\mathbb{N})$, that there exists a $i_k \in \{1, \ldots, n\}$ such that $k \in \sigma_{i_k}(\mathbb{N})$ and thus a $l_k \in \mathbb{N}$ such that $k = \sigma_{i_k}(l_k)$. Take now $N_m = \max(\{l_k | k \in \{1, \ldots, m\}\})$ then, as $\forall k \in \{1, \ldots, m\}$ we have $1 \leqslant l_k \leqslant N_m \Rightarrow l_k \in \{1, \ldots, N_m\}$, we have $k = \sigma_{i_k}(l_k) \in \sigma_{i_k}(\{1, \ldots, N_m\}) \subseteq \bigcup_{i \in \{1, \ldots, n\}} \sigma_i(\{1, \ldots, N_m\})$ proving that

$$\{1, \ldots, m\} \subseteq \bigcup_{i \in \{1, \ldots, n\}} \sigma_i(\{1, \ldots, N_m\}) \tag{17.52}$$

So using 17.100 we have

$$\begin{aligned} \sum_{i=1}^{m} x_i = \sum_{i \in \{1, \ldots, m\}} x_i &\leqslant \sum_{i \in \bigcup_{j \in \{1, \ldots, n\}} \sigma_j(\{1, \ldots, N_m\})} x_i \\ &\underset{17.50}{=} \sum_{i=1}^{n} \left(\sum_{j=1}^{N_m} x_{\sigma_i(j)} \right) \\ &\leqslant \sum_{i=1}^{n} \sup\left(\left\{ \left(\sum_{j=1}^{m} x_{\sigma_i(j)} \right) | m \in \mathbb{N} \right\} \right) \\ &= \sum_{i=1}^{n} \left(\sum_{j=1}^{\infty} x_{\sigma_i(j)} \right) \end{aligned}$$

which together with 17.51 proves the theorem. □

The following definition allows us to sum from $-\infty$ to $+\infty$

Definition 17.126. *Let* $\{x_i\}_{i \in \mathbb{Z}}$ *then* $\sum_{i=-\infty}^{\infty} x_i = \sum_{i=0}^{\infty} x_i + \sum_{i=1}^{\infty} x_{-i}$

17.2.4 Generalized series of positive extened reals

In the above section we had to introduce bijections and make distinctions between the finite case and the infinite case, to simplify notation we introduce the following definition.

Definition 17.127. *Let I be a countable set and* $\{x_i\}_{i \in I} \subseteq [0, \infty]$ *a family of extended positive reals then we define* $\sum_{i \in I} x_i$ *as follows, using the possible cases of I (see 5.30)*

- ***$I = \emptyset$.*** $\sum_{i \in I} x_i = 0$
- ***I* is finite non empty.** *then* $\sum_{\in I} x_i$ *is defined by 10.37*
- ***I* is infinite countable.** *then there exists a bijection* $\beta: \mathbb{N} \to I$ *and we define* $\sum_{i \in I} x_i = \sum_{i \in \mathbb{N}} x_{\beta(i)} = \sup(\{\sum_{i=1}^{n} x_i | n \in \mathbb{N}\})$ *(converges by 17.119)*

We must of course proof that this is well defined, more specifically for the infinite case (as the finite case is already proved in 10.37.

Proof. Let $\alpha: \mathbb{N} \rightarrow I$ be a another bijection then $\sigma: \mathbb{N} \rightarrow \mathbb{N}$ defined by $\sigma = \alpha \circ \beta^{-1}$ then we have $\sum_{i \in \mathbb{N}} x_{\beta(i)} \underset{17.119}{=} \sum_{i \in \mathbb{N}} x_{\sigma(\beta(i))} = \sum_{i \in \mathbb{N}} x_{\alpha(i)}$ □

As $\mathbb{1}_{\mathbb{N}}: \mathbb{N} \rightarrow \mathbb{N}$ is a bijection we have by the above definiton that

Example 17.128. if $\{x_i\}_{i \in \mathbb{N}}$ is a family of extended positive reals that $\sum_{i=1}^{\infty} x_i = \sum_{i \in \mathbb{N}} x_{\mathbb{1}_{\mathbb{N}}(i)} = \sum_{i \in \mathbb{N}} x_i$

We show now that we can always remove zeroes from the sum.

Theorem 17.129. *Let I be a countable set, $\{x_i\}_{i \in I}$ a family of extended positive reals and $J = \{i \in I | x_i \neq 0\}$ then $\sum_{i \in I} x_i = \sum_{i \in I} x_i$*

Proof. We have to consider three cases for I

***I* is finite.** Then as $I = I \setminus J \bigcup J$ and $I \bigcap (I \setminus J) = \emptyset$ we have $\sum_{i \in I} x_i \underset{10.45}{=} \sum_{i \in I \setminus J} x_i + \sum_{i \in J} x_i \underset{10.37}{=} 0 + \sum_{i \in I \setminus J} x_i = \sum_J x_i$

***I* is infinite countable.** Then there exists a bijection $\beta: \mathbb{N} \rightarrow I$ such that $\sum_{i \in I} x_i = \sum_{i=1}^{n} x_{\beta(i)}$ then we have

$$\begin{aligned} i \in \beta^{-1}(J) &\Leftrightarrow \beta(i) \in J \\ &\Leftrightarrow x_{\beta(i)} \neq 0 \\ &\Leftrightarrow i \in \{i \in \mathbb{N} | x_{\beta(i)} \neq 0 \end{aligned}$$

so that

$$\beta^{-1}(J) = \{i \in \mathbb{N} | x_{\beta(i)} \neq 0\} \tag{17.53}$$

then using 17.120 on $\{x_{\beta(i)}\}_{i \in \mathbb{N}}$ we have either that

$\boldsymbol{\beta^{-1}(J) = \emptyset}$**.** then $J = \emptyset$ [for if $j \in J$ then as β is surjective there exists a $i \in \mathbb{N}$ such that $\beta(i) = j \Rightarrow i \in \beta^{-1}(J) = \emptyset$ a contradiction] and $\sum_{i \in I} x_i = \sum_{i \in \mathbb{N}} x_{\beta(i)} = 0 \underset{\text{definition}}{=} \sum_{i \in \emptyset} x_i \underset{J = \emptyset}{=} \sum_{i \in J} x_i$.

$\boldsymbol{\emptyset \neq \beta^{-1}(J)}$ **is finite.** then there exists a $n \in \mathbb{N}$ and a bijection $\alpha: \{1, \ldots, n\} \rightarrow \beta^{-1}(J)$ such that

$$\sum_{i=1}^{\infty} x_{\beta(i)} = \sum_{i=1}^{n} x_{\beta(a\alpha(i))} \tag{17.54}$$

Further as $\beta \circ \alpha: \{1, \ldots, n\} \rightarrow \beta(\beta^{-1}(J)) \underset{\beta is a \text{bijection}}{=} J$ we have that

$$\begin{aligned} \sum_{i \in J} x_i &\underset{\text{def}}{=} \sum_{i \in \{1, \ldots, n\}} x_{(\beta \circ \alpha)(i)} \\ &= \sum_{i \in \{1, \ldots, n\}} x_{\beta(\alpha(i))} \\ &\underset{17.54}{=} \sum_{i=1}^{\infty} x_{\beta(i)} \\ &\underset{\text{def}}{=} \sum_{i \in I} x_i \end{aligned}$$

$\boldsymbol{\beta^{-1}(J)}$ is infinite countable. then there exists a bijection $\alpha: \mathbb{N} \rightarrow \beta^{-1}(J)$ such that

$$\sum_{i=1}^{\infty} x_{\beta(i)} = \sum_{i=1}^{\infty} x_{\beta(\alpha(i))} \tag{17.55}$$

Further as $\beta \circ \alpha: \mathbb{N} \rightarrow \beta(\beta^{-1}(J)) \underset{\beta \text{ is bijective}}{=} J$ we have that

$$\begin{aligned} \sum_{i \in J} x_i &\underset{\text{def}}{=} \sum_{i=1}^{\infty} x_{(\beta \circ \alpha)}(i) \\ &= \sum_{i=1}^{\infty} x_{\beta(\alpha(i))} \\ &\underset{17.55}{=} \sum_{i=1}^{\infty} x_{\beta(i)} \\ &\underset{\text{def}}{=} \sum_{i \in I} x_i \end{aligned}$$

So in all cases we have $\sum_{i \in I} x_i = \sum_{i \in J} x_i$. □

Corollary 17.130. *Let I be a countable set and $\{x_i\}_{i \in I}$ a set of extended positive reals, $J \subseteq I$ such that $\forall j \in J$ we have $x_j = \emptyset$ then $\sum_{i \in I} x_i = \sum_{i \in I \setminus J} x_i$*

Proof. First as $I \setminus J \subseteq I$ we have that $\{i \in I \setminus J | x_i \neq 0\} \subseteq \{i \in I | x_i \neq 0\}$. Second if $i \in \{i \in I | x_i\}$ then $i \in I \wedge x_i \neq 0$, if now $i \in J$ then $x_i = 0$ a contradiction so that $i \in I \setminus J$ proving that $i \in \{i \in I \setminus J | x_i \neq 0\}$. So we have that $\{i \in I | x_i \neq 0\} = \{i \in I \setminus J | x_i \neq 0\}$ hence $\sum_{i \in I} x_i \underset{17.129}{=} \sum_{i \in \{i \in I | x_i \neq 0\}} x_i = \sum_{i \in \{i \in I \setminus J | x_i \neq 0\}} x_i = \sum_{i \in I \setminus J} x_i$. □

Chapter 18

The exponential function

18.1 Power series

18.1.1 Definition and convergence domains

In this chapter we limit ourselves to series of real numbers so as $\langle\mathbb{R},+,0\rangle$ is a field we have that finite sums are always defined (see 10.2) So we can use the definition of a convergent (divergent) serie in $\mathbb{R}$ (see 12.363).

Theorem 18.1. *Let $\{x_i\}_{i\in\{k,\ldots,\infty\}}\subseteq\mathbb{R}$ then $\sum_{i=k}^{\infty} x_i$ converges if and only $\limsup_{n\rightarrow\infty}\left(\sum_{i=k}^{n} x_i\right)=\liminf_{n\rightarrow\infty}\left(\sum_{i=k}^{\infty} x_i\right)\in\mathbb{R}$ in which case we have $\sum_{i=k}^{\infty} x_i=\limsup_{n\rightarrow\infty}\left(\sum_{i=k}^{n} x_i\right)=\liminf_{n\rightarrow\infty}\left(\sum_{i=k}^{\infty} x_i\right)\in\mathbb{R}$*

Proof. This follows from 12.363 and 17.76 □

Definition 18.2. (Power Serie) *Let $z\in\mathbb{Z}$ and $\{a_n\}_{n\in\mathbb{N}_0}\subseteq\mathbb{C}$ then $\sum_{i=0}^{\infty} a_i\cdot z^i$ is called a power serie.*

Theorem 18.3. *Let $\{a_n\}_{n\in\mathbb{N}_0}\subseteq\mathbb{C}$ be a sequence then $a_0=\sum_{i=0}^{\infty} a_i\cdot 0^i$ and $|a_0|=\sum_{i=0}^{\infty}|a_i\cdot 0^i|=\sum_{i=0}^{\infty}|a_i|\cdot 0^i$. Let $n\in\mathbb{N}_0$ then $\sum_{i=1}^{n}|a_i\cdot 0^i|=\sum_{i=1}^{n}|a_i|\cdot|0|^i=|a_0|\cdot|0|^0+\sum_{i=1}^{\infty}|a_i|\cdot|0|^i=|a_0|+\sum_{i=1}^{\infty}|a_i|\cdot 0\cdot 0^{i-1}=|a_0|+0\cdot\sum_{i=1}^{\infty} a_i\cdot 0^{i-1}=|a_0|$.*

Proof. Let $n\in\mathbb{N}_0$ then $\sum_{i=1}^{n} a_i\cdot 0^i=a_0\cdot 0^0+\sum_{i=1}^{\infty} a_i\cdot o^i=a_0+\sum_{i=1}^{\infty} a_i\cdot 0\cdot 0^{i-1}=a_0+0\cdot\sum_{i=1}^{\infty} a_i\cdot 0^{i-1}=a_0$. □

Definition 18.4. *Let $\sum_{i=0}^{\infty} a_i\cdot z^i$ be a power serie then we define*

1. $\mathcal{D}_{\sum_{i=0}^{\infty} a_i\cdot z^i}=\{z\in\mathbb{C}|\sum_{i=0}^{\infty} a_i\cdot z^i\text{ is convergent}\}$, *$\mathcal{D}_{\sum_{i=0}^{\infty} a_i\cdot z^i}$ is called the **convergence domain** of $\sum_{i=0}^{\infty} a_i\cdot z^i$*

2. $\overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} = \{z \in \mathbb{C} | \sum_{i=1}^{\infty} a_i \text{ is absolute convergent}\}$, $\overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}}$ *is called the* ***absolute convergence domain*** *of* $\sum_{i=0}^{\infty} a_i \cdot z^i$

Theorem 18.5. *Let* $\sum_{i=0}^{\infty} a_i \cdot z^i$ *be a power serie then* $0 \in \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$

Proof. Using 18.3 we have that $0 \in \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}}$, further as absolute convergence implies convergence (see 12.382) we have that $\overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$ □

Theorem 18.6. *Let* $\sum_{i=0}^{\infty} a_i \cdot z^i$ *be a power serie then for* $0 \neq z_0 \in \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$ *we have* $\forall \rho$ *with* $0 \leqslant \rho < |z_0|$ *that* $\sum_{i=0}^{\infty} |a_n| \cdot \rho^n$ *is convergent.*

Proof. Let $0 \leqslant \rho < |z_0|$ then we have for δ either

$\rho = 0$. Using 18.3 we have that $\sum_{i=0}^{\infty} |a_n| \cdot \rho^n$ is convergent.

$0 < \rho$. As $z_0 \in \mathcal{D}_{\{a_i\}_{i \in \mathbb{N}_0}}$ we have that $\sum_{i=0}^{\infty} a_n \cdot z_0^i$ is convergent, so using 12.364 we have that $\lim_{n \to \infty} a_n \cdot z_0^n = 0$, hence by 12.318 $\{a_n \cdot z_0^n\}_{n \in \mathbb{N}}$ is Cauchy and thus bounded (see 12.324) proving

$$\exists K \in \mathbb{R} \text{ such that } \forall i \in \mathbb{N}_0 \vDash |a_n \cdot z^n| \leqslant K \tag{18.1}$$

Now $\forall n \in \mathbb{N}_0$ we have $|a_n| \cdot \rho^n \underset{0 \neq |z_0|}{=} |a_n| \cdot |z_0|^n \cdot \left(\frac{\rho}{|z_0|}\right)^n = |a_n z_0^n| \cdot q^n$ with $q = \left|\frac{\rho}{|z_0|}\right| < 1$ and using 18.1 we have then that

$$\forall n \in \mathbb{N}_0 \vDash |a_n| \cdot \rho^n \leqslant K \cdot q^n \text{ where } q < 1 \tag{18.2}$$

Using 12.403 we have that $\sum_{i=0}^{\infty} \rho^i$ is convergent and thus using 12.369 we have that $\sum_{i=0}^{\infty} K \cdot \rho^i$ is convergent. Using 18.2 on the above allows us to apply dominent convergence (see 12.381) to prove that $\sum_{i=0}^{\infty} |a_n| \cdot \rho^n$ is convergent. □

Corollary 18.7. *Let* $\sum_{i=0}^{\infty} a_i \cdot z^i$ *be a power serie,* $0 \neq z_0 \in \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$ *then* $B_{||}(0, |z_0|) \subseteq \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$

Proof. Take $z \in B_{||}(0, |z_0|)$ then for z we have either

$z = 0$. then we have $z = 0 \in \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}}$

$z \neq 0$. then $0 < |z| < |z_0| \Rightarrow \exists \rho \vdash 0 < |z| < \rho < |z_0|$. Using the previous theorem we have that $\sum_{i=0}^{\infty} |a_n| \cdot \rho^n$ is convergent. Finally as $\forall i \in \mathbb{N}_0$ we have $|a_i \cdot z^i| = |a_i| \cdot |z|^i < |a_i| \cdot \rho^i$ it follows from dominent vonvergence (see 12.381) that $\sum_{i=1}^{\infty} |a_i \cdot z^i|$ is convergent, hence $\sum_{i=0}^{\infty} a_i \cdot z^i$ is absolute convergent and thus $z \in \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}}$. □

From the above it follows that $0 \in [0, \infty[\bigcap \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \Rightarrow \emptyset \neq [0, \infty[\bigcap \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$ which allows us to define the following

Definition 18.8. *Let $\sum_{i=0}^{\infty} a_i \cdot z^i$ be a power serie then $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} = \sup \left(\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \bigcap [0, \infty[\right)$ is called the* ***conversion radius***

The above corollary is then summarized as follows

Corollary 18.9. *Let $\{a_i\}_{i \in \mathbb{N}_0} \subseteq \mathbb{C}$ be a sequence then we have for $R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$*

1. *If $0 \leqslant R_{\sum_{i=0}^{\infty} a_i \cdot z^i} < \infty$ then $B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i}) \subseteq \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \subseteq \overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})}$*
2. *If $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} = \infty$ then $\overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} = \mathbb{C} = \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$*
3. *If $\overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} = \mathbb{C} = \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$ then $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} = \infty$ and $B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i}) = \overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})}$*

Proof.

1. For $0 \leqslant R_{\sum_{i=0}^{\infty} a_i \cdot z^i} < \infty$ we have either

   $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} = 0$. if $0 \neq z \in \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$ then $0 < |z|$ so there exists a $\rho \in \mathbb{R}$ such that $0 < r < |z|$ proving that $0 < r \in [0, \infty[\bigcap B_{||}(0, |z|) \subseteq_{18.7} [0, \infty[\bigcap \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$. From this it follows that $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} = \sup \left(\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \bigcap [0, \infty[\right) > 0$ a contradiction. So we have that $B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i}) = \emptyset \subseteq \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} = \{0\} = \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \underset{0 \in \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}}{=} \{0\} = \overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})}$.

   $0 < R_{\sum_{i=0}^{\infty} a_i \cdot z^i} < \infty$. Let $z \in B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})$ then we have either

   z=0. then $z = 0 \in \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}}$

   $z \neq 0$. then $0 < |z| < R_{\sum_{i=0}^{\infty} a_i \cdot z^i} = \sup \left(\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \bigcap [0, \infty[\right)$ so there exists a $z' \in \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \bigcap [0, \infty[$ such that $0 < |z| < z' \leqslant R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$. Hence $z \in B_{||}(0, z') = B_{||}(0, |z'|) \subseteq_{18.7} \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}}$

   so in all cases we have $z \in B_{||}(0, R_{\{a_i\}_{i \in \mathbb{N}_0}})$ proving that

$$B_{||}(0, R_{\{a_i\}_{i \in \mathbb{N}_0}}) \subseteq \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \tag{18.3}$$

If $z \notin \overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})}$ then $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} < |z|$ then $\exists r \in [0, \infty[$ such that $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} < r < |z|$. If now $z \in \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$ then by 18.7 we have that $r \in \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \bigcap [0, \infty[$ hence $r \leqslant \sup \left(\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \bigcap [0, \infty[\right) = R_{\sum_{i=0}^{\infty} a_i \cdot z^i} < r$ a contradiction. So we must have that $\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \subseteq \overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})}$ proving using 18.3 that

$$B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i}) \subseteq \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \subseteq \overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})}$$

2. If $z \in \mathbb{C}$ then as $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} = \infty$ there exists by 17.11 a $r \in \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \bigcap [0, \infty[$ such that $|z| < r = |r| \Rightarrow z \in B_{||}(0, |r|) \subseteq_{18.7} \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}}$ proving that $\mathbb{C} \subseteq \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \subseteq \mathbb{C}$ giving

$$\overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} = \mathbb{C} = \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$$

3. As by (1) we have $B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i}) \subseteq \mathbb{C} \subseteq \overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})}$ we must have that $\overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})} = \mathbb{C}$. If now $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} < \infty$ then as $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} + 1 \in \mathbb{C} = \overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})}$ so that $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} < R_{\sum_{i=0}^{\infty} a_i \cdot z^i} + 1 \leqslant R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$ giving the contradiction $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} < R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$ so we must have that $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} = \infty$ which means by (2) that also $\mathbb{C} = \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$. □

We can simplify the above corollary if we extend the definition of a open/closed ball in $\mathbb{C}$ so that the radiuscan be infinite.

Definition 18.10. *Let $\delta \in [0, \infty]$ and $z_0 \in \mathbb{C}$ then $B_{||}(z_0, \delta) = \{z \in \mathbb{C} | |z - z_0| < \rho\}$ and $B_{||}(z_0, \delta) = \{z \in \mathbb{C} | |z - z_0| \leqslant \delta\}$ then we have*

1. *If $\delta = 0$ then $B_{||}(z_0, 0) = \emptyset$ a open set and $\overline{B_{||}(z_0, \delta)} = \{z_0\}$ a closed set*

2. *If $\delta < \infty$ then $B_{||}(z_0, \delta)$ is the classical open ball and $\overline{B_{||}(z_0, \delta)}$ is the classical closed ball.*

3. *If $\delta = \infty$ then if $z \in \mathbb{C}$ we have $|z - z_0| < |z - z_0| + 1 < \infty$ so that $B_{||}(z_0, \delta) = \mathbb{C} = \overline{B_{||}(z_0, \delta)}$ a open/closed set*

we have then the following

Lemma 18.11. *Let $\delta, \delta' \in [0, \infty]$ and $z_0 \in \mathbb{C}$ then if $B_{||}(z_0, \delta) \subseteq \overline{B_{||}(z_0, \delta')}$ then $\delta \leqslant \delta'$*

Proof. We have the following possibilities to consider for δ'

$\boldsymbol{\delta' \in]0, \infty[}$. assume then that $\delta' < \delta$ then there exists a $\rho \in \mathbb{R}$ such that $0 < \delta' < \rho < \delta$ hence $\delta' < |\rho| = \rho < \delta \Rightarrow \rho \in B_{||}(z_0, \delta) \Rightarrow \rho \in B_{||}(z_0, \delta') \Rightarrow \delta' < \rho \leqslant \delta' \Rightarrow \delta' < \delta'$ a contradiction so we must have that $\delta \leqslant \delta'$

$\delta'=0$. then $B_{||}(z_0,\delta)\subseteq\overline{B_{||}(z_0,\delta')}=\{0\}$ if now $0<\delta$ then there exists a $0<\rho<\delta\Rightarrow|\rho|=\rho<\delta\Rightarrow\rho\in B_{||}(z_0,\delta)\subseteq\{0\}\Rightarrow 0<\rho=0$ a contradiction so we must have $\delta\leqslant 0=\delta'$

$\delta'=\infty$. then $\delta\leqslant\infty=\delta'$

So in all cases we have $\delta\leqslant\delta'$. □

The above corollary condences then to

Corollary 18.12. *Let $\sum_{i=0}^{\infty} a_i\cdot z^i$ be a power serie then we have that $B_{||}(0,R_{\sum_{i=0}^{\infty}a_i\cdot z^i})\subseteq\overline{\mathcal{D}_{\sum_{i=0}^{\infty}a_i\cdot z^i}}\subseteq\mathcal{D}_{\sum_{i=0}^{\infty}a_i\cdot z^i}\subseteq\overline{B_{||}(0,R_{\sum_{i=0}^{\infty}a_i\cdot z^i})}$ where $B_{||}(0,R_{\sum_{i=0}^{\infty}a_i\cdot z^i})$ is open and $\overline{B_{||}(0,R_{\sum_{i=0}^{\infty}a_i\cdot z^i})}$ is closed.*

18.1.2 Convergence criteria

Lemma 18.13. *Let $n\in\mathbb{N}$ and $\{z_i\}_{i\in\{0,\ldots,n\}}\subseteq\mathbb{C}$, $\{\alpha_i\}_{i\in\{0,\ldots,n\}}\subseteq\mathbb{C}$ and $\{s_i\}_{i\in\{0,\ldots,n\}}$ defined by $s_i=\sum_{j=0}^{i} z_j$ then $\sum_{i=0}^{n}\alpha_i\cdot z_i=\sum_{i=0}^{n-1} s_i\cdot(\alpha_i-\alpha_{i+1})+\alpha_n\cdot s_n$*

Proof. We prove this by induction so let $\mathcal{S}=\{n\in\mathbb{N}|\text{if }\{z_i\}_{i\in\{0,\ldots,n\}}\subseteq\mathbb{C},\{\alpha_i\}_{i\in\{0,\ldots,n\}}\subseteq\mathbb{C}\text{ and }\{s_i\}_{i\in\{0,\ldots,n\}}\text{ defined by }s_i=\sum_{j=0}^{i}z_j\text{ then }\sum_{i=0}^{n}\alpha_i\cdot z_i=\sum_{i=0}^{n-1}s_i\cdot(\alpha_i-\alpha_{i+1})+\alpha_n\cdot s_n\}$ then we have

$1\in\mathcal{S}$. for

$$\begin{aligned}\sum_{i=0}^{n-1} s_i\cdot(\alpha_i-\alpha_{i+1})+\alpha_n\cdot s_n &\underset{n=1}{=} \sum_{i=0}^{0} s_i\cdot(\alpha_i-\alpha_{i+1})+\alpha_1\cdot s_1\\ &= s_0\cdot(\alpha_0-\alpha_1)+\alpha_1\cdot s_1\\ &= z_0\cdot(\alpha_0-\alpha_1)+\alpha_1\cdot(z_0+z_1)\\ &= \alpha_0\cdot z_0+\alpha_1\cdot z_1\\ &= \sum_{i=0}^{1}\alpha_i\cdot z_i\end{aligned}$$

proving that $1\in\mathcal{S}$

$n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$. We have

$$\begin{aligned}\sum_{i=0}^{n+1}\alpha_i\cdot z_i &= \left(\sum_{i=0}^{n}\alpha_i\cdot z_i\right)+\alpha_{n+1}\cdot z_{n+1}\\ &\underset{n\in\mathcal{S}}{=} \left(\sum_{i=0}^{n-1}s_i\cdot(\alpha_i-\alpha_{i+1})+\alpha_n\cdot s_n\right)+\alpha_{n+1}\cdot z_{n+1}\end{aligned}$$

$$\begin{aligned}
&= \left(\sum_{i=0}^{n-1} s_i\cdot(\alpha_i-\alpha_{i+1})\right)+(\alpha_n\cdot s_n-a_{n+1}\cdot s_n)+(\alpha_{n+1}\cdot z_{n+1}+\\
&\quad\alpha_{n+1}\cdot s_n)\\
&= \left(\sum_{i=0}^{n} s_i\cdot(\alpha_i-\alpha_{i+1})\right)+(\alpha_{n+1}\cdot z_{n+1}+\alpha_{n+1}\cdot s_n)\\
&= \left(\sum_{i=0}^{n} s_i\cdot(\alpha_i-\alpha_{i+1})\right)+\alpha_{n+1}\cdot\left(z_{n+1}+\sum_{i=0}^{n} z_i\right)\\
&= \left(\sum_{i=0}^{n} s_i\cdot(\alpha_i-\alpha_{i+1})\right)+\alpha_{n+1}\cdot s_{n+1}
\end{aligned}$$

proving that $n+1\in\mathcal{S}$ □

Lemma 18.14. (Inequality of Abel) *Let $n\in\mathbb{N}_0$ $\{z_i\}_{i\in\{0,\ldots,n\}}\subseteq\mathbb{C}$, $\{\alpha_i\}_{i\in\{0,\ldots,n\}}\subseteq[0,\infty[$ such that $\forall i\in\{0,\ldots,n-1\}$ we have $\alpha_{i+1}\leqslant\alpha_i$, $\sigma=\sum_{i=0}^{n}\alpha_i\cdot z_i$, $\{s_k\}_{k\in\{0,\ldots,n\}}$ defined by $s_k=\sum_{i=0}^{n} z_i$ and $\mu\in[0,\infty[$, then if $\forall k\in\{0,\ldots,n\}$ we have $|s_k|\leqslant\mu$ it follows that $|\sigma|\leqslant\alpha_0\cdot\mu$*

Proof. We have for n two cases to consider

$n=0$. then $|\sigma|=|\sum_{i=0}^{n}\alpha_i\cdot z_i|=|\alpha_0\cdot z_0|=|\alpha_0\cdot\sum_{i=0}^{0} z_i|=|\alpha_0\cdot s_0|=|\alpha_0|\cdot|s_0|\underset{0\leqslant\alpha_0}{=}\alpha_0\cdot|s_0|\leqslant\alpha_0\cdot\mu$ proving that $|\sigma|\leqslant\alpha_0\cdot\mu$

$n\neq 0$. then $n\in\mathbb{N}$ so we can use the previous lemma

$$\begin{aligned}
|\sigma| &= \left|\sum_{i=0}^{n}\alpha_i\cdot z_i\right|\\
&\underset{\text{previous lemma}}{=} \left|\sum_{i=0}^{n-1} s_i\cdot(\alpha_i-\alpha_{i+1})+\alpha_n\cdot s_n\right|\\
&\leqslant \sum_{i=0}^{n-1}|s_i\cdot(\alpha_i-\alpha_{i+1})|+|\alpha_n\cdot s_n|\\
&\underset{\alpha_{i+1}\leqslant\alpha_i,0\leqslant\alpha_n}{=} \sum_{i=0}^{n-1}|s_i|\cdot(\alpha_i-\alpha_{i+1})+\alpha_n\cdot s_n\\
&\leqslant \sum_{i=0}^{n-1}\mu\cdot(\alpha_i-\alpha_{i+1})+\alpha_n\cdot s_n\\
&= \mu\cdot\sum_{i=0}^{n-1}(\alpha_i-\alpha_{i+1})+\alpha_n\cdot\mu\\
&= \mu\cdot\left(\sum_{i=0}^{n-1}(\alpha_i-\alpha_{i+1})+\alpha_n\right)\\
&\underset{10.28}{=} \mu\cdot(\alpha_0-\alpha_n+\alpha_n)\\
&= \mu\cdot\alpha_0
\end{aligned}$$

which proves the theorem. □

The above lemma will be used to prove the convergence criteria of Abel Dirichlet

Theorem 18.15. (Abel Dirichlet) *Let $\{z_i\}_{i\in\mathbb{N}_0}\subseteq\mathbb{C}$, $\{s_n\}_{n\in\mathbb{N}_0}$ defined by $s_n=\sum_{i=0}^{n} z_i$, $\{\varepsilon_i\}_{i\in\mathbb{N}_0}\subseteq[0,\infty[$ such that $\forall i\in\mathbb{N}_0$ we have $\varepsilon_{i+1}\leqslant\varepsilon_i$ then we have*

1. *If $\exists C\in\mathbb{R}$ such that $\forall i\in\mathbb{N}_0$ we have $|s_i|\leqslant C$ and $\lim_{i\rightarrow\infty}\varepsilon_i=0$ then $\sum_{i=0}^{\infty}\varepsilon_i\cdot z_i$ is convergent*
2. *If $\sum_{i=0}^{\infty} z_i$ is convergent then $\sum_{i=0}^{\infty}\varepsilon_i\cdot z_i$ is convergent*

Proof.

1. Given $n,m\in\mathbb{N}\vdash n\leqslant m$ define $\{x_i^{(n)}\}_{i\in\{0,\ldots,m-n\}}$ by $x_i^{(n)}=z_{n+i}$, define $\{s_k^{(n)}\}_{k\in\{0,.,m-n\}}$ by $s_k^{(n)}=\sum_{i=0}^{k} x_i^{(n)}$ and define $\{\alpha_i\}_{i\in\{0,\ldots,m-n\}}$ by $\alpha_i^{(n)}=\varepsilon_{n+i}$ then we have that
$$\forall i\in\{0,\ldots,m-n-1\}\vDash\alpha_{i+1}^{(n)}\leqslant\alpha_i^{(m)} \text{ and } \{\{\alpha_i\}_{i\in\{0,\ldots,m-n\}}\}\subseteq[0,\infty[\tag{18.4}$$
Further giving $k\in\{0,\ldots,m-n\}$ we have $s_k^{(n)}=\sum_{i=0}^{k} x_i^{(n)}=\sum_{i=0}^{n+k} z_{n+i}=\sum_{i=n}^{n+k} z_i \underset{10.26}{=} \sum_{i=0}^{n+k} z_i-\sum_{i=0}^{n-1} z_i=s_{n+k}-s_{n-1}$ which in addition to $|s_{n+k}-s_{n-1}|\leqslant|s_{n+k}|+|s_n|\leqslant 2\cdot C$ proves that
$$\forall k\in\{0,\ldots,m-n\} \text{ we have } s_k^{(n)}=s_{n+k}-s_{n-1} \text{ and } \left|s_k^{(n)}\right|\leqslant 2\cdot C \tag{18.5}$$
Using the the ineaquality of Abele (see 18.14) we have then that $\left|\sum_{i=0}^{m-n}\alpha_i\cdot x_i^{(n)}\right|\leqslant\alpha_0\cdot(2\cdot C)$, and as $\sum_{i=0}^{m-n}\alpha_i\cdot x_i^{(n)}=\sum_{i=0}^{m-n}\varepsilon_{n+i}\cdot z_{n+i}=\sum_{i=n}^{m}\varepsilon_i\cdot z_i$ we have that
$$\left|\sum_{i=n}^{m}\varepsilon_i\cdot z_i\right|\leqslant 2\cdot C\cdot\varepsilon_n \tag{18.6}$$
Take now $\varepsilon>0$ then as $\lim_{n\rightarrow\infty}\varepsilon_i=0$ there exists a $N\in\mathbb{N}$ such that $\forall n\geqslant N$ we have that $\varepsilon_n=|\varepsilon_n-0|<\frac{\varepsilon}{2\cdot C}$. So we conclude using 18.6 that
$$\forall\varepsilon>0 \text{ there } \exists N\in\mathbb{N} \text{ such that } \left|\sum_{i=n}^{m}\varepsilon_i\cdot z_i\right|<\varepsilon$$
which by the Cauchy condition for series (see 12.365) and the fact that $\langle\mathbb{C},||\rangle$ proves that $\sum_{i=0}^{\infty}\varepsilon_i\cdot z_i$ converges.
2. From the hypothese it follows that $\{\varepsilon_i\}_{i\in\mathbb{N}_0}$ is a decreasing sequence, so using 17.66 we have that $\alpha=\lim_{i\rightarrow\infty}\varepsilon_i$ exists so that $\lim_{i\rightarrow\infty}(\varepsilon_i-\alpha)=0$. From (1) it follows then that $\sum_{i=1}^{n}(\varepsilon_i-\alpha)\cdot z_i$ converges. As by the hypothese $\sum_{i=0}^{\infty} z_i$ converges $\underset{12.369}{\Rightarrow}\sum_{i=0}^{\infty}\alpha\cdot z_i$ converges we have then by 12.369 that $\sum_{i=0}^{\infty}((\varepsilon_i-\alpha)\cdot z_i+\alpha\cdot z_i)$ converges proving that $\sum_{i=0}^{\infty}\varepsilon_i\cdot z_i$ □

Lemma 18.16. *Let $\{x_i\}_{i\in\{k,\ldots,\infty\}},\{y_i\}_{i\in\{k,\ldots,\infty\}}\subseteq]0,\infty[$ be such that $\forall i\in\{k,\ldots,\infty\}$ we have $\frac{x_{i+1}}{y_{i+1}}\leqslant\frac{x_i}{y_i}$ then if $\sum_{i=k}^{\infty} y_i$ is convergent then $\sum_{i=k}^{\infty} x_i$ is convergent.*

Proof. Take $\alpha = \frac{y_k}{x_k}$ then by the hypothese we have $\forall i \in \{k, \ldots, \infty\}$ that $\frac{x_k}{y_k} \leqslant \alpha \underset{0<y_k}{\Rightarrow} x_k \leqslant \alpha \cdot y_k \underset{0<x_k}{\Rightarrow} |x_k| \leqslant \alpha \cdot y_k$. As $\sum_{i=k}^{\infty} \alpha \cdot y_i$ is convergent (see 12.369) we have by the dominent convergence (see 12.381) that $\sum_{i=k}^{\infty} x_i$ is convergent. $\square$

Theorem 18.17. (d'Alembert) *Let* $\{x_i\}_{i \in \{k,.,\ldots,\infty\}} \subseteq]0, \infty[$ *and take* $\alpha = \liminf_{i \to \infty} \frac{x_{i+1}}{x_i}$ *and* $\beta = \limsup_{i \to \infty} \frac{x_{i+1}}{x_i}$ *then we have*

1. *If* $\beta < 1$ *then* $\sum_{i=k}^{\infty} x_i$ *is convergent*

2. *If* $1 < \alpha$ *then* $\sum_{i=k}^{\infty} x_i$ *is divergent (meaning* $\sum_{i=k}^{\infty} x_i$ *is not convergent)*

Proof.

1. As $\beta < 1$ there exists a ρ such that $0 \leqslant \beta < \rho < 1$ then we have $\inf\left(\left\{\sup\left(\left\{\frac{x_{i+1}}{x_i} | i \in \{n, \ldots, \infty\}\right\}\right) | n \in \{k, \ldots, \infty\}\right\}\right) < \rho$ so there exists a $n \in \{k, \ldots, \infty\}$ such that $\sup\left(\left\{\frac{x_{i+1}}{x_i} | i \in \{n, \ldots, \infty\}\right\}\right) < \rho$ hence
$$\forall i \in \{n, \ldots, \infty\} \text{ that } \frac{x_{i+1}}{x_i} < \rho \underset{0<\rho}{=} \frac{\rho^{i+1}}{\rho^i} \tag{18.7}$$
Now as $\rho < 1$ we have by 12.403 that $\sum_{i=1}^{\infty} \rho^i$ is convergent and thus using 12.367 we have
$$\sum_{i=n}^{\infty} \rho^i \tag{18.8}$$
Using lemma 18.16 we have by 18.7 and 18.8 that $\sum_{i=n}^{\infty} x_i$ is convergent. Applying 12.367 again we have that $\sum_{i=k}^{\infty} x_i$ is convergent.

2. If $\alpha < 1$ then we have $1 < \sup\left(\left\{\inf\left(\left\{\frac{x_{i+1}}{x_i} | i \in \{n, \ldots, \infty\}\right\}\right) | n \in \{k, \ldots, \infty\}\right\}\right)$ hence there exists a $n \in \{k, \ldots, \infty\}$ such that $1 < \inf\left(\left\{\frac{x_{i+1}}{x_i} | i \in \{n, \ldots, \infty\}\right\}\right)$ hence $\forall i \in \{n, \ldots, \infty\}$ we have that $1 < \frac{x_{i+1}}{x_i} \Rightarrow x_i < x_{i+1}$ so $\{x_i\}_{i \in \{n, \ldots, \infty\}}$ is a increasing sentence. Using 17.83 we have that $\lim_{n \to \infty} x_i = \sup(\{x_i | i \in \{k, \ldots, \infty\}\}) \geqslant x_k > 0 \Rightarrow \lim_{i \to \infty} x_i \neq 0$. If now $\sum_{k=1}^{\infty} x_i$ is convergent then we have by 12.364 that $\lim_{i \to \infty} x_i = 0$ a contradiction proving that $\sum_{k=1}^{\infty} x_i$ is divergent. $\square$

Corollary 18.18. *Let* $z \in \mathbb{C}$ *then* $\sum_{n=0}^{\infty} \frac{|z|^n}{n!}$ *and* $\sum_{n=0}^{\infty} \frac{z^n}{n!}$ *converges and* $\sum_{n=0}^{\infty} \frac{0^i}{i!} = 1$. *So we have* $B_{||}\left(0, R_{\sum_{i=0}^{\infty} \frac{z^i}{i!}}\right) = \overline{B_{||}\left(0, R_{\sum_{i=0}^{\infty} \frac{z^i}{i!}}\right)} = \mathcal{D}_{\sum_{n=0}^{\infty} \frac{z^n}{n!}} = \mathbb{C} = \overline{\mathcal{D}_{\sum_{n=0}^{\infty} \frac{z^n}{n!}}}$

Proof. Let $z \in \mathbb{C}$ then we have either

$\boldsymbol{z = 0}$. then using 18.3 we have that $\sum_{n=0}^{\infty} \frac{|z|^n}{n!}$ and $\sum_{n=0}^{\infty} \frac{z^n}{n!}$ converges

$\boldsymbol{z \neq 0}$. then we have $\forall n \in \{0, \ldots, \infty\}$ that $0 < \frac{|z|^n}{n}$ and $\frac{\frac{|z|^{n+1}}{(n+1)!}}{\frac{|z|^n}{n!}} = \frac{|z|}{n+1}$. Using the consequence of the Archimedian property of the reals (see 8.62) there exists a $k \in \{0, \ldots, \infty\}$ such that $2 \cdot |z| < k$ so that $\forall n \geqslant k$ we have $|z| < k \Rightarrow \frac{|z|}{k} < \frac{1}{2}$ proving that $\sup\left(\left\{\frac{\frac{|z|^{n+1}}{(n+1)!}}{\frac{|z|^n}{n!}} | n \in \{k, \ldots, \infty\}\right\}\right) \leqslant \frac{1}{2}$ hence $\inf\left(\left\{\sup\left(\left\{\frac{\frac{|z|^{n+1}}{(n+1)!}}{\frac{|z|^n}{n!}} | n \in \{k, \ldots, \infty\}\right\}\right) | n \in \{0, \ldots, \infty\}\right\}\right) \leqslant \inf\left(\left\{\sup\left(\left\{\frac{\frac{|z|^{n+1}}{(n+1)!}}{\frac{|z|^n}{n!}} | n \in \{k, \ldots, \infty\}\right\}\right) | n \in \{0, \ldots, \infty\}\right\}\right) \leqslant \frac{1}{2} < 1$. Applying then d'Alembert (see 18.17) we conclude that $\sum_{i=0}^{\infty} \frac{|z|^i}{i!}$ converges. Using dominent convergence (see 12.381) we have then that $\sum_{n=0}^{\infty} \frac{z^n}{n!}$ is convergent. □

Using the above corollary the following definition is well defined

Definition 18.19. (Exponential Function) $\exp: \mathbb{C} \to \mathbb{C}$ *is defined by* $z \to \sum_{n=0}^{\infty} \frac{z^n}{n!}$

18.1.3 Differentiation of power series

Definition 18.20. *Let* $\{a_i\}_{i \in \mathbb{N}_0}$ *be a sequence and* $D_{\sum_{i=0}^{\infty} a_i \cdot z^i}$ *its convergence domain then we can define* $\sum_{n=0}^{\infty} a_n (.)^n : D_{\{a_i\}_{i \in \mathbb{N}_0}} \to \mathbb{C}$ *by* $z \to (\sum_{n=0}^{\infty} a_n \cdot (.)^n)(z) = \sum_{n=0}^{\infty} a_n \cdot z^n$

If $0 < R_{\{a_i\}_{i \in \mathbb{N}_0}}$ then $0 \in B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i}) \subseteq D_{\{a_i\}_{i \in \mathbb{N}_0}}$ so that it makes sense to talk about the derivative of $\sum_{n=0}^{\infty} a_n (.)^n$ at 0 on $B_{||}(0, R_{\{a_i\}_{i \in \mathbb{N}_0}})$. If we would differentiate the different terms of $\sum_{i=0}^{\infty} a_n \cdot z^n$ in a informal way we would have $\sum_{i=0}^{\infty} n \cdot a_n \cdot z^{n-1} = \sum_{i=1}^{\infty} n \cdot a_n \cdot z^{n-1} = \sum_{i=0}^{\infty} (n+1) \cdot a_{n+1} \cdot z^n$. Of course we must find out that $\sum_{i=0}^{\infty} (n+1) \cdot a_{n+1} \cdot z^n$ is convergent and what it's convergence domain is before even proving that this is the derivative of $\sum_{i=0}^{\infty} a_n \cdot z^n$.

Definition 18.21. *Let* $\sum_{i=0}^{\infty} a_i \cdot z^i$ *be a power serie then* $\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i$ *is called the* ***derivated power serie of*** $\sum_{i=0}^{\infty} a_i \cdot z^i$

Lemma 18.22. *Let* $\sum_{i=0}^{\infty} a_i \cdot z^i$ *be a powerserie then*

1. $\mathcal{D}_{\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$
2. *If* $0 < R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$ *then* $R_{\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i} = R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$

Proof.

1. *Let* $z \in \mathcal{D}_{\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i}$ *then* $\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i$ *is convergent. Now for the sequence* $\left\{\frac{1}{i+1}\right\}_{i \in \mathbb{N}_0}$ *we have that* $\forall i \in \mathbb{N}_0$ *that* $\frac{1}{(i+1)+1} < \frac{1}{(i+1)}$ *so using Abel Dirichlet (see 18.15) we have that* $\sum_{i=0}^{\infty} a_{i+1} \cdot z^i = \sum_{i=0}^{\infty} \frac{i+1}{i+1} \cdot a_{i+1} \cdot z^i$ *is convergent. Using 12.369 we have then that* $\sum_{i=0}^{\infty} a_{i+1} \cdot z^{i+1} = \sum_{i=0}^{\infty} z \cdot a_{i+1} \cdot z^i$ *is convergent. As* $a_0 \cdot z^0 + \sum_{i=0}^{\infty} a_{i+1} \cdot z^{i+1} \underset{12.366}{=} a_0 \cdot z^0 + \sum_{i=1}^{\infty} a_i \cdot z^i \underset{12.367}{=} \sum_{i=0}^{\infty} a_i \cdot z^i$ *we have that* $\sum_{i=0}^{\infty} a_i \cdot z^i$ *converges proving that* $z \in \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$

2. *First as* $\overline{B_{||}(0, R_{\sum_{i=0}^{\infty}(i+1)\cdot a_{i+1}\cdot z^i})} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty}(i+1)\cdot a_{i+1}\cdot z^i} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i\cdot z^i} \subseteq \overline{B_{||}(0, R_{\sum_{i=0}^{\infty}\cdot a_i\cdot z^i})}$ *so using 18.11 we have that*
$$R_{\sum_{i=0}^{\infty}(i+1)\cdot a_{i+1}\cdot z^i} \leqslant R_{\sum_{i=0}^{\infty} a_i\cdot z^i} \tag{18.9}$$
Let $0 < z \in \mathbb{C}$ *be such that* $|z| < R_{\sum_{i=0}^{\infty} a_i\cdot z^i}$ *then* $\exists \rho \in \mathbb{R}$ *such that* $0 < |z| < \rho < R_{\sum_{i=0}^{\infty} a_i\cdot z^i}$*. From* $\rho < R_{\sum_{i=0}^{\infty} a_i\cdot z^i}$ *it follows that* $\sum_{i=0}^{\infty} a_n \cdot \rho^n$ *converges. Using 12.364 we have then that* $\exists K \in \mathbb{R}_+$ *such that*
$$\forall n \in \mathbb{N}_0 \vDash |a_n| \cdot \rho^n = |a_n \cdot \rho^n| < K \tag{18.10}$$
Now $\forall n \in \mathbb{N}_0$ *we have*
$$\begin{aligned} |(n+1)\cdot a_{n+1}\cdot z^n| &= (n+1)\cdot|a_{n+1}|\cdot|z|^n \\ &= \frac{n+1}{\rho}\cdot\left(\frac{|z|}{\rho}\right)^n\cdot a_{n+1}\cdot\rho^{n+1} \\ &\leqslant_{18.10} K\cdot\frac{n+1}{\rho}\cdot\left(\frac{|z|}{\rho}\right)^n \\ &= K\cdot\frac{n+1}{\rho}\cdot q^n \text{ where } 0 < q = \frac{|z|}{\rho} < 1 \end{aligned}$$
proving that
$$\forall n \in \mathbb{N}_0 \vDash |(n+1)\cdot a_{n+1}\cdot z^n| \leqslant K\cdot\frac{n+1}{\rho}\cdot q^n \text{ where } 0 < q < 1 \tag{18.11}$$
Now for the serie $\sum_{n=0}^{\infty}(n+1)\cdot q^n$ *we have that* $\frac{((n+1)+1)\cdot q^{n+1}}{(n+1)\cdot q^n} = \frac{n+2}{n+1}\cdot q$*. As by 12.321 we have that* $\lim_{n\to\infty}\frac{n+2}{n+1} = 1 \Rightarrow \lim_{n\to\infty}\left(\frac{n+2}{n+1}\cdot q\right) = q < 1$ *we can use d'Alembert (see 18.17) proving to conclude that* $\sum_{i=0}^{\infty}(n+1)\cdot q^n$ *is convergent. Using then dominent convergence (see 12.381) together with 18.11 we conclude then that* $\sum_{i=0}^{\infty}(n+1)\cdot a_{n+1}\cdot z^n$ *is convergent. So we have proved that* $\forall x \in \mathbb{C}$ *with* $0 < |z| < R_{\sum_{i=0}^{\infty} a_i\cdot z^i}$ *we have* $z \in \mathcal{D}_{\sum_{i=0}^{\infty}(i+1)\cdot a_{i+1}\cdot z^i}$ *and as* $0 \in \mathcal{D}_{\sum_{i=0}^{\infty}(i+1)\cdot a_{i+1}\cdot z^i}$ *we conclude that* $B_{||}(0, R_{\sum_{i=0}^{\infty} a_i\cdot z^i}) \subseteq \mathcal{D}_{\sum_{i=0}^{\infty}(i+1)\cdot a_{i+1}\cdot z^i} \subseteq \overline{B_{||}(0, R_{\sum_{i=0}^{\infty}(i+1)\cdot a_{i+1}\cdot z^i})}$*. From 18.11 it follows then that* $R_{\sum_{i=0}^{\infty} a_i\cdot z^i} \leqslant R_{\sum_{i=0}^{\infty}(i+1)\cdot a_{i+1}\cdot z^i}$ *which together with 18.9 proves finally that*
$$R_{\sum_{i=0}^{\infty}(i+1)\cdot a_{i+1}\cdot z^i} = R_{\sum_{i=0}^{\infty} a_i\cdot z^i}$$
□

We now prove that the derivated power serie of a power serie is indeed its derivative.

Theorem 18.23. *Let* $\sum_{i=0}^{\infty} a_i \cdot z^i$ *be a power serie with* $0 < R_{\sum_{i=0}^{\infty} a_i\cdot z^i}$ *and* $z \in \mathbb{C}$ *then* $\sum_{i=0}^{\infty} a_i\cdot(.)^i : B_{||}(0, R_{\sum_{i=0}^{\infty} a_i\cdot z^i}) \to \mathbb{C}$ *has a derivate (see 14.16) at every* $z \in B_{||}(0, R_{\sum_{i=0}^{\infty} a_i\cdot z^i})$ *with* $(\sum_{i=0}^{\infty} a_i\cdot(.)^i)'(z) = (\sum_{i=0}^{\infty}(i+1)\cdot a_{i+1}\cdot(.)^i)(z)$

Proof. Define $f = \sum_{i=0}^{\infty} a_i\cdot(.)^i$ and $g = \sum_{i=0}^{\infty}(i+1)\cdot a_{i+1}\cdot(.)^i$. Let $\varepsilon > 0$ and take z such that $|z| < R_{\sum_{i=0}^{\infty} a_i\cdot z^i} \Rightarrow z \in B_{||}(0, R_{\sum_{i=0}^{\infty} a_i\cdot z^i})$. As $B_{||}(0, R_{\sum_{i=0}^{\infty} a_i\cdot z^i})$ is open there exists a $0 < \rho < R_{\sum_{i=0}^{\infty} a_i\cdot z^i}$ such that $B_{||}(z, \rho) \subseteq B_{||}(0, R_{\sum_{i=0}^{\infty} a_i\cdot z^i})$, if now $h \in \mathbb{C}$ is choosen such that $|h| < \min(\rho, \rho - |z|)$ then $|h + z| \leqslant |h| + |z| < \rho < R_{\sum_{i=0}^{\infty} a_i\cdot z^i}$ so that $f(z)$, $f(z+h)$ and $g(z)$ are well defined (see also 18.22) and $|h+z-z|=|h|<\rho$ so that $h\in B_{||}(z,\rho)$ (needed for the definition of a derivative)

$$
\begin{aligned}
\frac{f(z+h)-f(z)}{h}-g(z) &= \frac{1}{h}\cdot(f(z+h)-f(z)-h\cdot g(z))\\
&= \frac{1}{h}\cdot\left(\sum_{i=0}^{\infty} a_i\cdot(z+h)^i-\sum_{i=0}^{\infty} a_i\cdot z^i-h\cdot\sum_{i=0}^{\infty}(i+1)\cdot a_{i+1}\cdot\right.\\
&\quad\left. z^i\right)\\
&= \frac{1}{h}\cdot\left(\sum_{i=0}^{\infty} a_i\cdot\left(\sum_{j=0}^{i}\binom{i}{j}\cdot h^j\cdot z^{i-j}\right)-\sum_{i=0}^{\infty} a_i\cdot z^i-\right.\\
&\quad\left. h\cdot\sum_{i=0}^{\infty}(i+1)\cdot a_{i+1}\cdot z^i\right)\\
&= \frac{1}{h}\cdot\left(\sum_{i=0}^{\infty} a_i\cdot\left(\left(\sum_{j=0}^{i}\binom{i}{j}\cdot h^j\cdot z^{i-j}\right)-z^i\right)-h\cdot\right.\\
&\quad\left.\sum_{i=0}^{\infty}(i+1)\cdot a_{i+1}\cdot z^i\right)\\
&= \frac{1}{h}\cdot\left(\sum_{i=1}^{\infty} a_i\cdot\left(\left(\sum_{j=0}^{i}\binom{i}{j}\cdot h^j\cdot z^{i-j}\right)-z^i\right)+a_0\cdot\right.\\
&\quad\left(\left(\sum_{j=0}^{0}\binom{i}{j}\cdot h^j\cdot z^{i-j}\right)-z^0\right)-h\cdot\sum_{i=0}^{\infty}(i+1)\cdot a_{i+1}\cdot\\
&\quad\left. z^i\right)\\
&= \frac{1}{h}\cdot\left(\sum_{i=1}^{\infty} a_i\cdot\left(\left(\sum_{j=0}^{i}\binom{i}{j}\cdot h^j\cdot z^{i-j}\right)-z^i\right)+a_0\cdot\right.\\
&\quad\left.\left(\binom{0}{0}\cdot h^0\cdot z^0-1\right)-h\cdot\sum_{i=0}^{\infty}(i+1)\cdot a_{i+1}\cdot z^i\right)\\
&= \frac{1}{h}\cdot\left(\sum_{i=1}^{\infty} a_i\cdot\left(\left(\sum_{j=0}^{i}\binom{i}{j}\cdot h^j\cdot z^{i-j}\right)-z^i\right)-h\cdot\right.\\
&\quad\left.\sum_{i=0}^{\infty}(i+1)\cdot a_{i+1}\cdot z^i\right)
\end{aligned}
$$

$$= \frac{1}{h} \cdot \left(\sum_{i=1}^{\infty} a_i \cdot \left(\left(\left(\sum_{j=1}^{i} \binom{i}{j} \cdot h^j \cdot z^{i-j} + \binom{i}{0} \cdot h^0 \cdot z^{i-0} \right) - z^i \right) - h \cdot \sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i \right)$$

$$= \frac{1}{h} \cdot \left(\sum_{i=1}^{\infty} a_i \cdot \left(\left(\left(\sum_{j=1}^{i} \binom{i}{j} \cdot h^j \cdot z^{i-j} + z^i \right) - z^i \right) - h \cdot \sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i \right)$$

$$= \frac{1}{h} \cdot \left(\sum_{i=1}^{\infty} a_i \cdot \left(\sum_{j=1}^{i} \binom{i}{j} \cdot h^j \cdot z^{i-j} \right) - h \cdot \sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i \right)$$

$$= \frac{1}{h} \cdot \left(\sum_{i=2}^{\infty} a_i \cdot \left(\sum_{j=1}^{i} \binom{i}{j} \cdot h^j \cdot z^{i-j} \right) + a_1 \cdot \left(\sum_{j=1}^{1} \binom{i}{j} \cdot h^j \cdot z^{i-j} \right) - h \cdot \sum_{i=1}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i - h \cdot (0+1) \cdot a_{0+1} \cdot z^0 \right)$$

$$= \frac{1}{h} \cdot \left(\sum_{i=2}^{\infty} a_i \cdot \left(\sum_{j=1}^{i} \binom{i}{j} \cdot h^j \cdot z^{i-j} \right) + a_1 \cdot \binom{1}{1} \cdot h^1 \cdot z^{1-1} - h \cdot \sum_{i=1}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i - h \cdot a_1 \cdot z^0 \right)$$

$$= \frac{1}{h} \cdot \left(\sum_{i=2}^{\infty} a_i \cdot \left(\sum_{j=1}^{i} \binom{i}{j} \cdot h^j \cdot z^{i-j} \right) - h \cdot \sum_{i=1}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i \right)$$

$$
\begin{aligned}
&\underset{12.366}{=} \frac{1}{h}\cdot\left(\sum_{i=2}^{\infty} a_i\cdot\left(\sum_{j=1}^{i}\left(\begin{array}{c} i\\ j\end{array}\right)\cdot h^j\cdot z^{i-j}\right)-h\cdot\sum_{i=2}^{\infty} i\cdot a_i\cdot z^{i-1}\right)\\
&= \frac{1}{h}\cdot\left(\sum_{i=2}^{\infty} a_i\cdot\left(\left(\sum_{j=1}^{i}\left(\begin{array}{c} i\\ j\end{array}\right)\cdot h^j\cdot z^{i-j}\right)-i\cdot h\cdot z^{i-1}\right)\right)\\
&= \sum_{i=2}^{\infty} a_i\cdot\left(\left(\sum_{j=1}^{i}\left(\begin{array}{c} i\\ j\end{array}\right)\cdot h^{j-1}\cdot z^{i-j}\right)-i\cdot z^{i-1}\right)\\
&= \sum_{i=2}^{\infty} a_i\cdot\left(\left(\sum_{j=2}^{i}\left(\begin{array}{c} i\\ j\end{array}\right)\cdot h^{j-1}\cdot z^{i-j}\right)+\left(\begin{array}{c} i\\ 1\end{array}\right)\cdot h^{1-1}\cdot z^{i-1}-i\cdot z^{i-1}\right)\\
&= \sum_{i=2}^{\infty} a_i\cdot\left(\left(\sum_{j=2}^{i}\left(\begin{array}{c} i\\ j\end{array}\right)\cdot h^{j-1}\cdot z^{i-j}\right)+i\cdot z^{i-1}-i\cdot z^{i-1}\right)\\
&= \sum_{i=2}^{\infty} a_i\cdot\left(\sum_{j=2}^{i}\left(\begin{array}{c} i\\ j\end{array}\right)\cdot h^{j-1}\cdot z^{i-j}\right)
\end{aligned}
$$

to summarize

$$
\frac{f(z+h)-f(z)}{h}-h\cdot g(z)=\sum_{i=2}^{\infty} a_i\cdot\left(\sum_{j=2}^{i}\left(\begin{array}{c} i\\ j\end{array}\right)\cdot h^{j-1}\cdot z^{i-j}\right) \tag{18.12}
$$

Now let $i\in\{2,\ldots,\infty\}$ then we have

$$
\begin{aligned}
\left|\sum_{j=2}^{i}\left(\begin{array}{c} i\\ j\end{array}\right)\cdot h^{j-1}\cdot z^{i-j}\right| &= |h|\cdot\left|\sum_{j=2}^{i}\left(\begin{array}{c} i\\ j\end{array}\right)\cdot h^{j-2}\cdot z^{i-j}\right|\\
&\leqslant |h|\cdot\left|\sum_{j=2}^{i}\left(\begin{array}{c} i\\ j\end{array}\right)\cdot |h^{j-2}|\cdot|z^{i-j}|\right|\\
&= |h|\cdot\left|\sum_{j=2}^{i}\frac{i!}{j!\cdot(i-j)!}\cdot|h^{j-2}|\cdot|z^{i-j}|\right|
\end{aligned}
$$

$$
\begin{aligned}
&\underset{i,j\geqslant 2}{=} && |h|\cdot\left|\sum_{j=2}^{i}\frac{i\cdot(i-1)\cdot(i-2)!}{j\cdot(j-1)\cdot(j-2)!\cdot(i-j)!}\cdot |h^{j-2}|\cdot|z^{i-j}|\right| \\
&\leqslant_{j\geqslant 2\Rightarrow j\cdot(j-1)\geqslant 1} && |h|\cdot\sum_{j=2}^{i}\frac{i\cdot(i-1)\cdot(i-2)!}{(j-2)!\cdot(i-j)!}\cdot|h^{j-2}|\cdot|z^{i-j}| \\
&= && |h|\cdot i\cdot(i-1)\cdot\sum_{j=2}^{i}\frac{(i-2)!}{(j-2)!\cdot((i-2)-(j-2))!}\cdot|h^{j-2}|\cdot |z^{(i-2)-(j-2)}| \\
&= && |h|\cdot i\cdot(i-1)\cdot\sum_{j=0}^{i-2}\frac{(i-2)!}{j!\cdot((i-2)-j)!}\cdot|h^{j}|\cdot |z^{(i-2)-j}| \\
&= && |h|\cdot i\cdot(i-1)\cdot(|z|+|h|)^{i-2}
\end{aligned}
$$

Using the above with 18.12 gives

$$\left|\frac{f(z+h)-f(z)}{h}-g(z)\right|\leqslant |h|\cdot\sum_{i=2}^{\infty}(i\cdot(i-1)\cdot a_i\cdot(|z|+|h|)^{i-2}) \tag{18.13}$$

Now as $|z|+|h|<\rho<R_{\sum_{i=0}^{\infty}a_i\cdot z^i}$ we have using 18.6 that $\sum_{i=0}^{\infty}a_i\cdot(|z|+|h|)^i$ converges. From 18.22 it follows then that $\sum_{i=0}^{\infty}(i+1)\cdot a_{i+1}\cdot(|z|+|h|)^i$ converges and applying 18.22 we have that $\sum_{i=0}^{\infty}(i+1)\cdot(i+2)\cdot a_{i+2}\cdot(|z|+|h|)^i$ converges. As $\sum_{i=0}^{\infty}(i+1)\cdot(i+2)\cdot a_{i+2}\cdot(|z|+|h|)^i \underset{12.366}{=} \sum_{i=2}^{\infty}(i-1)\cdot i\cdot a_i\cdot(|z|+|h|)^{i-2}$ proving that

$$0<\sum_{i=2}^{\infty}(i\cdot(i-1)\cdot a_i\cdot(|z|+|h|)^{i-2})\in\mathbb{R} \tag{18.14}$$

If we now add the extra condition on h that $0<|h|<\min\left(\rho,\rho-|z|,\frac{\varepsilon}{\sum_{i=2}^{\infty}(i\cdot(i-1)\cdot a_i\cdot(|z|+|h|)^{i-2})}\right)$ then by 18.3 we have

$$\left|\frac{f(z+h)-f(z)}{h}-g(z)\right|<\varepsilon$$

which proves our theorem. □

We can now prove what the derivative of the exponential function is.

Corollary 18.24. *The exponential function* $\exp:\mathbb{C}\rightarrow\mathbb{C}$ *defined by* $\exp(z)=\sum_{n=0}^{\infty}\frac{z^n}{n!}$ *has a derivative equal to itself at every point in* $\mathbb{C}$, *in other words* $\forall z\in\mathbb{C}$ *that* $\exp'(z)=\exp(z)$

Proof. First by 18.18 we have that $\mathcal{D}_{\sum_{n=0}^{\infty}\frac{z^n}{n!}} = \mathbb{C}$. So if $z \in C$ we have that using the above theorem that $\exp'(z)$ exists and that $\exp'(z) = \sum_{i=0}^{\infty} (i+1) \cdot \frac{z^i}{(i+1)!} = \sum_{i=0}^{\infty} \frac{x^i}{i!} = \exp(z)$ □

Corollary 18.25. *The exponential function* $\exp: \mathbb{C} \to \mathbb{C}$ *defined by* $\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$ *is continuous on* $\mathbb{C}$

Proof. Let $z \in \mathbb{C}$ then by the previous corollary exp has a derivate at z and thus exp is differentiable at z (see 14.17). As differentiability implies continuity (see 14.10) we have that exp is continuous at z hence exp is continuous on $\mathbb{C}$. □

Next we show that the exponential function behaves as a power which motivates the other notation for $\exp(z)$ as e^z.

Theorem 18.26. *Let* $x, y \in \mathbb{C}$ *then* $\exp(x+y) = \exp(x) \cdot \exp(y)$

Proof. Let $x, y \in \mathbb{C}$ then as $\mathcal{D}_{\sum_{n=0}^{\infty}\frac{z^n}{n!}} = \mathbb{C} = \overline{\mathcal{D}_{\sum_{n=0}^{\infty}\frac{z^n}{n!}}}$ we have that $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ and $\exp(y) = \sum_{n=0}^{\infty} \frac{y^n}{n!}$ are absolute convergent

$$\begin{aligned}
\exp(x) \cdot \exp(y) &= \left(\sum_{n=0}^{\infty} \frac{x^n}{n!}\right) \cdot \left(\sum_{n=0}^{\infty} \frac{y^n}{n!}\right) \\
&\underset{12.390}{=} \left(\sum_{n \in \mathbb{N}_0} \frac{x^n}{n!}\right) \cdot \left(\sum_{n \in \mathbb{N}_0} \frac{y^n}{n!}\right) \\
&\underset{12.399}{=} \sum_{(i,j) \in \mathbb{N}_0 \times \mathbb{N}_0} \left(\frac{x^i}{i!} \cdot \frac{y^j}{j!}\right) \{\text{absolute convergent}\}
\end{aligned}$$

giving

$$\exp(x) \cdot \exp(y) = \sum_{(i,j) \in \mathbb{N}_0 \times \mathbb{N}_0} \left(\frac{x^i}{i!} \cdot \frac{y^j}{j!}\right) \text{a absolute convergent serie} \tag{18.15}$$

Let $n \in \mathbb{N}_0$ define then $N_n = \{(i,j) \in \mathbb{N}_0 \times \mathbb{N}_0 | i + j = n\}$, as $N_n \subseteq \mathbb{N}_0 \times \mathbb{N}_0$ we have $\bigcup_{n \in \mathbb{N}_0} N_n \subseteq \mathbb{N}_0 \times \mathbb{N}_0$ and if $(i,j) \in \mathbb{N}_0 \times \mathbb{N}_0$ then either $(i,j) \in N_{i+j}$ proving that $\mathbb{N}_0 \times \mathbb{N}_0 \subseteq \bigcup_{n \in \mathbb{N}_0} N_n$. Further we have that if $n.n' \in \mathbb{N}_0$ with $n = n'$ then if $(i,j) \in N_n \bigcap N_{n'} \Rightarrow i + j = n \wedge i + j = n' \Rightarrow n = n'$ a contradiction so we must have $N_n \bigcap N_{n'} = \emptyset$. Summarized

$$\mathbb{N}_0 \times \mathbb{N}_0 = \bigsqcup_{i \in \mathbb{N}_0} N_n \tag{18.16}$$

Further we have that $\beta: \{0, \ldots, n\} \to N_n$ defined by $\beta(i) = (i, n-i)$ is a bijection

injectivity. if $\beta(i) = \beta(j)$ then $(i, n-i) = (j, n-j) \Rightarrow i = j$

surjectivity. if $(i,j) \in N_n$ then $i + j = n \Rightarrow j = n - i \Rightarrow (i,j) = (i, n-i) = \beta(i)$

Using then the associativity of absolute convergent series we have

$$\begin{aligned}
\sum_{(i,j)\in\mathbb{N}_0\times\mathbb{N}_0}\left(\frac{x^i}{i!}\cdot\frac{y^j}{j!}\right) &\underset{12.398}{=} \sum_{n\in\mathbb{N}_0}\left(\sum_{(i,j)\in N_n}\left(\frac{x^i}{i!}\cdot\frac{y^j}{j!}\right)\right)\\
&\underset{12.391}{=} \sum_{n\in\mathbb{N}_0}\left(\sum_{i\in\{0,\ldots,n\}}\left(\frac{x^{\beta(i)_1}\cdot y^{\beta(i)_2}}{\beta(i)_1!\cdot\beta(j)_2!}\right)\right)\\
&= \sum_{n\in\mathbb{N}_0}\left(\sum_{i\in\{0,\ldots,n\}}\left(\frac{x^i\cdot y^{(n-i)}}{i!\cdot j!}\right)\right)\\
&= \sum_{n\in\mathbb{N}_0}\frac{1}{n!}\left(\sum_{i\in\{0,\ldots,n\}}\left(\frac{n!\cdot x^i\cdot y^{(n-i)}}{i!\cdot n-i!}\right)\right)\\
&= \sum_{n\in\mathbb{N}_0}\frac{1}{n!}\left(\sum_{i\in\{0,\ldots,n\}}\left(\left(\begin{array}{c}n\\i\end{array}\right)\cdot x^i\cdot y^{(n-i)}\right)\right)\\
&= \sum_{n\in\mathbb{N}_0}\frac{(x+y)^i}{n!}\\
&= \exp(x+y)
\end{aligned}$$

which by 18.15 proves that

$$\exp(x+y)=\exp(x)\cdot\exp(y)$$ □

Theorem 18.27. $\exp\colon\mathbb{C}\to\mathbb{C}$ *satisfies*

1. $\exp(0)=1$
2. $\forall z\in\mathbb{C}$ *we have* $\exp(z)\neq 0$
3. $\forall z\in\mathbb{C}$ *we have* $\exp(-z)=\frac{1}{\exp(z)}$
4. $\forall z\in\mathbb{R}$ *we have* $0<\exp(z)$ *in other words* $\exp(\mathbb{R})=\mathbb{R}_+$
5. $\forall z\in\mathbb{R}$ *with* $x>0$ *we have* $\exp(x)>1$
6. $\forall x,y\in\mathbb{R}$ *with* $x>y$ *we have* $\exp(x)>\exp(y)$ *[*$\exp$ *is strictly increasing on* $\mathbb{R}$*]*

Proof.

1. $\exp(0)=\sum_{i=0}^{\infty}\frac{z^i}{i!}\underset{18.3}{=}\frac{z^0}{0!}=1$
2. Assume that $\exists z\in\mathbb{C}$ such that $\exp(z)=0$ then as $1=\exp(0)=\exp(z+(-z))\underset{18.26}{=}\exp(z)\cdot\exp(-z)=0\cdot\exp(-z)=0$ leading to the contradiction $1=0$. Hence we have $\forall z\in\mathbb{C}$ that $\exp(z)\neq 0$
3. $1=\exp(0)=\exp(z+(-z))\underset{18.26}{=}\exp(z)\cdot\exp(-z)\underset{\exp(z)\neq 0}{\Rightarrow}\exp(-z)=\frac{1}{\exp(z)}$

4. If $z \in \mathbb{R}$ then we have 3 cases to consider for z

   $z \in \mathbb{R}_+$. then $\forall i \in \mathbb{N}_0$ we have $0 < \frac{z^i}{i!}$ (see 9.68) so that by 12.371 $0 < \sum_{i=0}^{\infty} \frac{x^i}{i!} = \exp(z)$

   $z = 0$. then $0 < 1 = \exp(z)$

   $z \in \mathbb{R}_-$. then $-z \in \mathbb{R}_+$ so that $0 < \frac{1}{\exp(z)} = \exp(-z)$

5. If $z > 0$ then by 9.68 $\forall i \in \mathbb{N}$ we have $0 < z^i \Rightarrow 0 < \frac{z^i}{i!}$ so that using 12.371

$$0 < \sum_{i=1}^{\infty} \frac{z^i}{i!} \tag{18.17}$$

   So $\exp(z) = \sum_{i=0}^{\infty} \frac{z^i}{i!} = \frac{z^0}{0!} + \sum_{i=1}^{\infty} \frac{z^i}{i!} = 1 + \sum_{i=1}^{\infty} \frac{z^i}{i!} > 1$

6. Let $x > y$ then $x - y > 0$ so by 18.27(5) we have $\exp(x) = \exp(y + (x - y)) = \exp(y) \cdot \exp(x - y) > \exp(y)$ □

Corollary 18.28. *Let $x, y \in \mathbb{R}$ with $x < y$ then* exp *is convex on* $[x, y]$

Proof. This follows from the fact that $\forall x \in \mathbb{R}$ we have that $\exp''(x) = (\exp'(x))' \underset{18.24}{=} \exp'(x) \underset{18.24}{=} \exp(x) > 0$ and the condition for convexitivity (see 14.94). □

Definition 18.29. (Euler's number) $e = \exp(1)$

We can now see how for natural numbers exp behaves as a power of e

Theorem 18.30.

1. $\forall n \in \mathbb{N}_0$ *we have* $\exp(n) = e^n$
2. $\forall n \in \mathbb{N}_0$ *we have* $\exp(-n) = \frac{1}{e^n} \underset{\text{definition}}{=} e^{-n}$

Proof.

1. *We use mathematical induction so let $\mathcal{S} = \{n \in \mathbb{N}_0 | \exp(n) = e^n\}$ then we have*

   $0 \in \mathcal{S}$. $\exp(0) = 1 = e^0$

   $n \in S \Rightarrow n + 1 \in \mathcal{S}$. *then* $\exp(n+1) \underset{18.26}{=} \exp(n) \cdot \exp(1) \underset{n \in \mathcal{S}}{=} e^n \cdot e = e^{n+1}$ *proving that* $n + 1 \in \mathcal{S}$

2. $\exp(-n) \underset{18.27}{=} \frac{1}{\exp(n)} \underset{(1)}{=} \frac{1}{e^n} e^{-n}$ □

Theorem 18.31. *Let $y \in \mathbb{R}_+$ then there exists a $x \in \mathbb{R}$ such that $y < \exp(x)$*

Proof. As $0 < 1$ we have by 18.27 that $1 = \exp(0) < \exp(1) = e$. Using 9.64 there exists a $n \in \mathbb{N}$ such that $y < e^n \underset{\text{previous theorem}}{=} \exp(n)$ □

Next we want to define the Napierian logarithm which is the inverse of exp resticted to $\mathbb{R}$, so we must prove that $\exp_{|\mathbb{R}}: \mathbb{R} \to \mathbb{R}_+$ is a bijection.

Theorem 18.32. $\exp_{|\mathbb{R}|}: \mathbb{R} \to \mathbb{R}_+$ *is a bijection where* $\exp_{|\mathbb{R}|}(\mathbb{R}_-) =]0,1[$ *and* $\exp_{|\mathbb{R}}(\mathbb{R}_+) =]1,\infty[$

Proof.

injectivity. Let $\exp(x) = \exp(y)$ and assume that $x \neq y$ then we may assume that $x > y$, as exp is strictly increasing (see 18.27 (6)) we have $\exp(x) > \exp(y)$ a contradiction so we must have that $x = y$.

surjectivity. Let $y \in \mathbb{R}_+ =]0,\infty[$ then we must consider 3 cases

$y = 1$. then $y = 1 = \exp(0)$

$1 < y$. then by 18.31 there exists a $b \in \mathbb{R}$ such that $y < \exp(b)$ so that we have $\exp(0) < y < \exp(b)$. As exp is continuous on $\mathbb{R}$ (see 18.25) and thus on $[0,b]$ we have by the itermediate value theorem (see 12.444) that $\exists x \in [0,b]$ such that $\exp(x) = y$

$0 < y < 1$. So $1 < \frac{1}{y}$ and we can apply the previous case to find a $x \in \mathbb{R}$ such that $\exp(x) = \frac{1}{y}$ hence $\exp(-x) = \frac{1}{\exp(x)} = y$

for the remaing of the theorem we have if $x \in \mathbb{R}_+$ by 18.27 (5) that $1 < \exp(x) \Rightarrow \exp(x) \in]1,\infty[$ so that

$$\exp(\mathbb{R}_+) \subseteq]1,\infty[\tag{18.18}$$

If $x \in \mathbb{R}_-$ then $-x \in \mathbb{R}_+$ so that $1 < \exp(-x) = \frac{1}{\exp(x)}$ or $1 < \frac{1}{\exp(x)} \Rightarrow \exp(x) < 1$ proving that $\exp(x) \in]0,1[$ so that

$$\exp(\mathbb{R}_-) \subseteq]0,1[\tag{18.19}$$

If now $y \in]1,\infty[\subseteq \mathbb{R}_+$ then as exp is a surjection there exists a $x \in \mathbb{R}$ such that $\exp(x) = y$ assume $x \in \mathbb{R}_-$ then by 18.19 we have $y = \exp(x) \in]0,1[$ not $]1,\infty[$ so we must have $x \in \mathbb{R}_+$ proving that $]1,\infty[\subseteq \exp(\mathbb{R}_+)$ which together with 18.18 gives

$$\exp(\mathbb{R}_+) =]1,\infty[$$

If now $y \in]0,1[\subseteq \mathbb{R}_+$ then as exp is a surjection there exists a $x \in \mathbb{R}$ such that $\exp(x) = y$ assume $x \in \mathbb{R}_+$ then by 18.18 we have $y = \exp(x) \in]1,\infty[$ not $]0,1[$ so we must have $x \in \mathbb{R}_-$ proving that $]0,1[\subseteq \exp(\mathbb{R}_-)$ which together with 18.19 gives

$$\exp(\mathbb{R}_-) =]0,1[$$

□

The above definition allows us to define the Napierian logarithm

Definition 18.33. (Napierian logarithm) $\log: \mathbb{R}_+ \to \mathbb{R}$ *is defined by* $\log = \exp^{-1}$

Based on the properties of exp we have the following properties of the logarithme

Theorem 18.34. $\log: \mathbb{R}_+ \to \mathbb{R}$ *has the following properties*

1. $\forall x \in \mathbb{R}_+$ *we have that* $\exp(\log(x)) = x$

2. *$\forall x \in \mathbb{R}$ we have that* $\log(\exp(x)) = x$
3. $\log(e) = 1$
4. $\log(1) = 0$
5. *$\forall x, y \in \mathbb{R}_+$ we have* $\log(x \cdot y) = \log(x) + \log(y)$
6. *$\forall x \in \mathbb{R}_+$ we have* $\log\left(\frac{1}{x}\right) = -\log(x)$
7. *$\forall x \in \mathbb{R}_+$* $\log$ *has a derivative* $\log'(x) = \frac{1}{x}$
8. $\log: \mathbb{R}_+ \rightarrow \mathbb{R}$ *is strictly increasing and $\forall x < 1$ we have* $\log(x) < 0$

Proof. This follows easely from the definition of log as the inverse of exp on $\mathbb{R}_+$

1. Let $x \in \mathbb{R}$ then $\exp(\log(x)) = \exp(\exp^{-1}(x)) = x$
2. Let $x \in \mathbb{R}$ then $\log(\exp(x)) = \exp^{-1}(\exp(x)) = x$
3. $\log(e) = \log(\exp(1)) \underset{(2)}{=} 1$
4. As $\exp(0) = 1$ we have that $\log(1) = \exp^{-1}(\exp(0)) = 0$
5. Let $x, y \in \mathbb{R}_+$ then $\exp(\log(x) + \log(y)) = \exp(\log(x)) \cdot \exp(\log(y)) \underset{(1)}{=} x \cdot y$ so that $\log(x \cdot y) = \log(\exp(\log(x) + \log(y))) \underset{(2)}{=} \log(x) + \log(y)$
6. Let $x \in \mathbb{R}_+$ then $x = \exp(\log(x))$ so that $\exp(-\log(x)) = \frac{1}{x}$ and thus $-\log(x) = \log(\exp(-\log(x))) = \log\left(\frac{1}{x}\right)$
7. Take $x \in \mathbb{R}_+$ and take $y = \log(x)$. As $\exp(y) > 0$ we have by a consequence of the inverse function theorem (see 14.155) that there exists a open $U \subseteq \mathbb{R}_+$, $V \subseteq \mathbb{R}$ with $y \in U$ and $x = \exp(\log(x)) = \exp(y) \in V$ such that $\exp_{|U}: U \rightarrow V$ has a inverse $\exp_{|U}^{-1}: V \rightarrow U$ such that $\forall z \in V$ we have $(\exp_{|U}^{-1})'(z) \cdot \exp(\exp^{-1}(z))$. As $\forall w \in U$ we have that $(\exp_{|U} \circ \log_{|V})(w) = \exp_{|U}(\log_{|V}(w)) = \exp(\log(w)) = w = 1_V(w)$ and $\forall w \in U$ we have that $(\log_{|V} \circ \exp_{|U})(w) = \log(\exp(w)) = w = 1_U(w)$ we conclude that $\exp_{|U}^{-1} = \log_{|V}$. So we have that $\forall z \in V$ that $1 = (\log_{|V})'(z) \cdot \exp(\log(z)) \underset{14.18}{=} \log'(z) \cdot z$. More specific as $x \in V \subseteq \mathbb{R}_+$ we proved that $\log'(x) = \frac{1}{x}$
8. Let $x, y \in \mathbb{R}_+$ with $x < y$ then if $\log(y) \leqslant \log(x)$ we have if exp is stritly increasing we have $y = \exp(\log(y)) \leqslant \exp(\log(x)) = x$ contradicting $x < y$ so that $\log(x) < \log(y)$. Further if $x < 1$ then $\log(x) < \log(1) = 0$. □

Once we have introduced the log function and the definition $e^x = \exp(x)$ (which was expired by the fact that $\exp(n) = e^n = \underbrace{e \ldots e}_{n \,\text{times}}$ (see 18.30) we can now define a general definiton of the power of a real number to a real number.

Definition 18.35. *Let $a \in]0, \infty[$ be a positive number and $x \in \mathbb{R}$ then* $a^{[x]} = \exp(x \cdot \log(a)) \in \mathbb{R}$

Remark 18.36. The restrictions in the above definition are needed for $\log(a)$ is only defined if $a \in \mathbb{R}_+ =]0, \infty[$, further if for example if $x = \frac{1}{2}$ then we will see that $a^{\frac{1}{2}} = \sqrt{a}$ which is not defined in the reals if $a = \frac{1}{2}$. Later we will see that in the case of $0 < x$ we can include the case $a = 0$ in the definition of a^x.

Next we prove now that we have the classical properties of our generalized power.

Theorem 18.37. *Let $a \in]0, \infty[= \mathbb{R}_+$ then we that*

1. *$\forall n \in \mathbb{N}_0$ we have that $a^{[n]} = \underbrace{a \ldots a}_{n \text{ times}}$ (we have a conflict of notation here so we use $\underbrace{a \ldots a}$ to note the classical power on $\mathbb{R}$)*
2. *$\forall x \in \mathbb{R}$ we have $1^{[x]} = 1$*
3. *$\forall x \in \mathbb{R}$ we have $e^{[x]} = \exp(x)$*
4. *$\forall x \in \mathbb{R}$ we have $\frac{a^{[x]}}{a} = a^{[x-1]}$*
5. *$\forall x, y \in \mathbb{R}$ we have $(a^{[x]})^{[y]} = a^{[x \cdot y]}$*
6. *$\forall x \in \mathbb{R}$ we have $a^{[-x]} = \frac{1}{a^{[x]}}$*
7. *$\forall b \in]0, \infty[$ and $\forall x \in \mathbb{R}$ we have $(a \cdot b)^{[x]} = a^{[x]} \cdot b^{[x]}$*
8. *$\forall x \in \mathbb{R}$ we have $\log(a^{[x]}) = x \cdot \log(a)$*
9. *Let $a \in \mathbb{R}_+$ and define $a^{(\cdot)}: \mathbb{R} \to \mathbb{R}_+$ by $a^{(\cdot)}(x) = a^{[x]}$ then $\forall x \in \mathbb{R}$ we have $(a^{(\cdot)})'(x) = \log(a) \cdot a^{[x]}$*
10. *Let $x \in \mathbb{R}$ and define $(.)^{[x]}: \mathbb{R}_+ \to \mathbb{R}_+$ by $(.)^{[x]}(y) = y^{[x]}$ then $\forall y \in \mathbb{R}$ we have that $((.)^{[x]})'(y) = x \cdot y^{[x-1]}$*

Proof.

1. We proof this by induction so let $\mathcal{S} = \left\{ n \in \mathbb{N}_0 | a^{[n]} = \underbrace{a \ldots a}_{n \text{ times}} \right\}$ then we have

 $\mathbf{0} \in \boldsymbol{\mathcal{S}}$. $a^{[0]} = \exp(0 \cdot \log(a)) = \exp(0) \underset{18.27}{=} 1 = \underbrace{a \ldots a}_{0 \text{ times}}$ proving that $n \in \mathbb{N}_0$

 $\boldsymbol{n} \in \boldsymbol{\mathcal{S}} \Rightarrow \boldsymbol{n} + \mathbf{1} \in \boldsymbol{\mathcal{S}}$. then

$$\begin{array}{rcl} a^{[n+1]} & = & \exp((n+1) \cdot \log(a)) \\ & = & \exp(n \cdot \log(a) + \log(a)) \\ & \underset{18.26}{=} & \exp(n \cdot \log(a)) \cdot \exp(\log(a)) \\ & \underset{n \in \mathcal{S}}{=} & \underbrace{a \ldots a}_{n \text{ times}} \cdot \exp(\log(a)) \\ & = & \underbrace{a \ldots a}_{n \text{ times}} \cdot a \\ & = & \underbrace{a \ldots a}_{n+1 \text{ times}} \end{array}$$

proving that $n+1\in\mathcal{S}$

2. Let $x\in\mathbb{R}$ then $x^{[1]}=\exp(\log(1)\cdot x) \underset{18.34\ (4)}{=} \exp(0\cdot x)=\exp(0) \underset{18.27}{=} 1$

3. $e^{[x]}=\exp(x\cdot\log(e)) \underset{18.34\ (3)}{=} \exp(x\cdot 1)=\exp(x)$

4.
$$\begin{aligned}(a^{[x]})^{[y]} &= (\exp(x\cdot\log(a)))^{[y]}\\ &= \exp(y\cdot\log(\exp(x\cdot\log(a))))\\ &= \exp(y\cdot(x\cdot\log(a)))\\ &= \exp((x\cdot y)\cdot\log(a))\\ &= a^{[x]\cdot[y]}\end{aligned}$$

5.
$$\begin{aligned}\frac{a^{[x]}}{a} &= \frac{\exp(\log(a)\cdot x)}{\exp(\log(a))}\\ &\underset{18.27\ (3)}{=} \exp(-\log(a))\cdot\exp(\log(a)\cdot x)\\ &= \exp(\log(a)\cdot x-\log(a))\\ &= \exp(\log(a)\cdot(x-1))\\ &= a^{[x-1]}\end{aligned}$$

6. Let $x\in\mathbb{R}$ then $a^{[-x]}=\exp(-x\cdot\log(a)) \underset{18.27\ (3)}{=} \frac{1}{\exp(x\cdot\log(a))}=\frac{1}{a^{[x]}}$

7. Let $b\in]0,\infty[$ and $x\in\mathbb{R}$ then $(a\cdot b)^{[x]}=\exp(\log(a\cdot b)\cdot x) \underset{18.34\ (5)}{=} \exp((\log(a)+\log(b))\cdot x)=\exp(\log(a)\cdot x+\log(b)\cdot x) \underset{18.26}{=} \exp(\log(a)\cdot x)\cdot\exp(\log(b)\cdot x)=a^{[x]}\cdot b^{[x]}$

8. $\log(a^{[x]})=\log(\exp(x\cdot\log(a)))=x\cdot\log(a)$

9. If we define $f:\mathbb{R}\to\mathbb{R}$ by $f(x)=\log(a)\cdot x$ then we have that $(\exp\circ f)(x)=\exp(f(x))=\exp(\log(a))=a^{[x]}=(a^{(\cdot)})(x)$ proving that $\exp\circ f=a^{(\cdot)}$. As $\forall x\in\mathbb{R}$ we have that $\exp'(x)$ and $f'(x)$ exists we have by 14.29 that $(a^{(')})'(x)=\exp'(f(x))\cdot f'(x)=\exp(f(x))\cdot\log(a)=\log(a)\cdot\exp(\log(a)\cdot x)=\log(a)\cdot a^x$

10. Let $x\in\mathbb{R}$ and define $f:\mathbb{R}_+\to\mathbb{R}$ by $f(y)=\log(y)\cdot x$ then $\forall y\in\mathbb{R}$ we have $f'(y)=x\cdot\log(y) \underset{18.34\ (7)}{=} \frac{x}{y}$. Further as $(\exp\circ f)(y)=\exp(f(y))=\exp(\log(y)\cdot x)=a^{[x]}=((.)^x)(y)$ we conclude that $(.)^x=\exp\circ f$. So using 14.29 we have $\forall x\in\mathbb{R}$ that $((.)^{[x]})'(y)=\exp(f(y))\cdot f'(y)=\exp(f(y))\cdot\frac{x}{y}=x\cdot\frac{\exp(\log(y)\cdot x)}{y}=x\cdot\frac{y^{[x]}}{x} \underset{(4)}{=} x\cdot y^{[x-1]}$ $\square$

Using the above theorem (9) we see that the function $(.)^x:\mathbb{R}_+\to\mathbb{R}_+$ has a derivative and thus by 14.17 and 14.10 we have that $(.)^{[x]}$ is continuous at every $x\in\mathbb{R}_+=]0,\infty[$ this can be extended to $[0,\infty[$ using the subspace topology on $[0,\infty[$ if we extend the definition of a power to $[0,\infty[$ which can be done if we restrict the power to strict positive numbers.

Definition 18.38. *Let* $p \in]0, \infty[$ *and* $a \in [0, \infty[$ *then we define* $a^p = \begin{cases} a^{[p]} = \exp(p \cdot \log(a)) \text{ if } a \in]0, \infty[\\ 0 \text{ if } a = 0 \end{cases}$

Note that the requirement $0 < p$ is needed for if $a \in]0, \infty[$ then $a^0 = \exp(\log(a) \cdot 0) = 1$ and as $0^0 = 0$ the function $[.]^0$: $[0, \infty[\rightarrow \mathbb{R}$ defined by $[.]^0(x)$ is then discontinuous at 0. Furter if $p < 0$ then $0^p = 0^{-p} = \frac{1}{0^p}$ is not well defined. We have now simular properties for a^p as a^p

Theorem 18.39. *Let* $a \in [0, \infty[= \mathbb{R}_+$ *then we that*

1. $\forall p \in]0, \infty[$ *we have* $a^p = 0 \Leftrightarrow a = 0$
2. $\forall n \in \mathbb{N}$ *we have that* $a^n = \underbrace{a \ldots a}_{n \, \text{times}}$ *(we have a conflict of notation here so we use* $\underbrace{a \ldots a}$ *to note the classical power on* $\mathbb{R}$*)*
3. $\forall x \in]0, \infty[= \mathbb{R}_+$ *we have* $1^x = 1$
4. $\forall x \in]0, \infty[$ $e^x = \exp(x)$
5. $\forall x \in]1, \infty[$ *and* $a \neq 0$ *then* $\frac{a^x}{a} = a^{x-1}$
6. $\forall x, y \in]0, \infty[$ *we have* $(a^x)^y = a^{x \cdot y}$
7. $\forall b \in [0, \infty[$ *and* $\forall x \in \mathbb{R}$ *we have* $(a \cdot b)^x = a^x \cdot b^x$
8. $\forall p \in]0, \infty[$ *we have that* $[.]^p$: $[0, \infty] \rightarrow \mathbb{R}$ *defined by* $[.]^p(x) = x^p$ *is strictly increasing*

Proof.

1. *If* $a^p = 0$ *if* $a \neq 0$ *we have* $a^p = \exp(p \cdot \log(a)) > 0$ *a contradiction so we must have* $a = 0$. *The opposite equivalence follows from* $0^p = 0$
2. *If* $n \in \mathbb{N}$ *then if* $a \in]0, \infty[$ *we have* $a^n = a^{[n]} \underset{18.37 \, (1)}{=} \underbrace{a \ldots a}_{n \, \text{times}}$ *and* $0^n = 0 = \underbrace{0 \ldots 0}_{1 \, \text{time}}$
3. *Let* $x \in]0, \infty[$ *then as* $1 \in]0, \infty[$ *we have that* $1^x = 1^{[x]} \underset{18.37 \, (2)}{=} 1$
4. *Let* $x \in]0, \infty[$ *then as* $e \neq 0$ *we have* $e^x = e^{[x]} \underset{18.37 \, (3)}{=} \exp(x)$
5. *Let* $x \in]1, \infty[$ *and* $a \neq 0$ *then* $\frac{a^x}{a} = \frac{a^{[x]}}{a} \underset{18.37}{=} a^{[x-1]} \underset{0 < x-1}{=} a^{x-1}$
6. $\forall x, y \in]0, \infty[$ *we have for* $a \in [a, \infty[$ *that*

   $\boldsymbol{a = 0.}$ *then* $(0^x)^y = 0^y = 0 = 0^{x \cdot y}$

   $\boldsymbol{a \neq 0.}$ *then* $a^x = a^{[x]} \neq 0$ *(see (1)) so that* $(a^x)^y = (a^{[x]})^{[y]} \underset{18.37}{=} a^{[x \cdot y]} = a^{x \cdot y}$
7. *We have for* a, b *either*

   $\boldsymbol{a = 0 \wedge b = 0.}$ *then* $(a \cdot b)^x = 0^x = 0 = 0 \cdot 0 = 0^x \cdot 0^x = a^x \cdot b^x$

$a=0\wedge b\neq 0$. *then* $(a\cdot b)^x=0^x=0\cdot b^x=0^x\cdot b^x=a^x\cdot b^x$

$a\neq 0\wedge b=0$. *then* $(a\cdot b)^x=0^x=a^x\cdot 0=a^x\cdot 0^x=a^x\cdot b^x$

$a\neq 0\wedge b\neq 0$. $(a\cdot b)^x \underset{a\cdot b\neq 0}{=} (a\cdot b)^{[x\cdot y]} \underset{18.37\ (7)}{=} =a^{[x]}\cdot b^{[x]}=a^x\cdot b^x$

8. *Let* $x,y\in[0,\infty[$ *with* $x<y$ *then we have for* x *either*

$x=0$. *then* $0<y$ *and* $[.]^p(x)=[.]^p(0)=0<\exp(p\cdot\log(x))=[.]^p(x)$

$0<x$. *then* $\log(x)<_{18.34\ (8)} \underset{0<p}{\Rightarrow} p\cdot\log(x)<p\cdot\log(y) \underset{18.27(6)}{\Rightarrow} \exp(p\cdot\log(x))<\exp(p*\log(y))$ *so that* $[.]^p(x)=\exp(p\cdot\log(x))<\exp(p\cdot\log(y))=[.]^p(y)$ □

Theorem 18.40. *Let* $p\in]0,\infty[$ *then* $[.]^p\colon[0,\infty[\rightarrow\mathbb{R}$ *defined by* $[.]^p(x)=x^p$ *then we have* $[.]^p$ *is continuous* $\forall x\in[0,\infty[$ *[where we use the canonical topology on* $\mathbb{R}$ *based on the norm* $||$ *and the subspace topology on* $[0,\infty[$ *(see 12.6)*

Proof. Let $x\in[0,\infty[$ the we can divide the proof in two pieces

$x\neq 0$. Let $[.]^p(x)\in B_{||}(x,\varepsilon)$ then as $(.)^p\colon]0,\infty[\rightarrow\mathbb{R}$ is continuous $(((.)^p)'(x)$ exists together with 14.17 and 14.10) there exists a $B_{||}(x,\delta_1)$ such that $\forall y\in B_{||}(x,\delta_1)\bigcap]0,\infty[$ we have $[.]^p(y)=y^p=(.)^p(y)\in B_{||}(x,\varepsilon)$. Further as $x\neq 0$ we have $0<x$ take then $\delta=\min(x,\delta_1)$. If $0\in B_{||}(x,\delta)$ then $x=x-0<\delta\leqslant x$ a contradiction so that $0\notin B_{||}(x,\delta)$ which proves that $B_{||}(x,\delta)\bigcap]0,\infty[=B_{||}(x,\delta)\bigcap[0,\infty[$. Finally $\forall y\in B_{||}(x,\delta)\bigcap[0,\infty[=B_{||}(x,\delta)\bigcap]0,\infty[\subseteq B_{||}(x,\delta_1)\bigcap]0,\infty[$ we have that $[.]^p(y)\in B_{||}(x,\varepsilon)$ proving that $[.]^p$ is continuous at x.

$x=0$. Let $\varepsilon>0$ then as $\exp\colon\mathbb{R}\rightarrow\mathbb{R}_+$ is a bijection there exists a $\delta_1\in\mathbb{R}$ such that $\exp(\delta_1)=\varepsilon$. Further as $\log\colon\mathbb{R}^+\rightarrow\mathbb{R}$ is a bijection there exists a $\delta<\mathbb{R}_+$ such that $\log(\delta)=\frac{\delta_1}{p}$. Hence if $y\in B_{||}(0,\delta)\bigcap[0,\infty[$ then $y<\delta$ so that $\log(y)<\log(\delta)$ [as log is strictly increasing], as $0<p$ if follows that $p\cdot\log(y)<p\cdot\log(\delta) \underset{\text{exp is increasing}}{\Rightarrow} \exp(p\cdot\log(y))<\exp(p\cdot\log(\delta))=\exp\left(p\cdot\frac{\delta_1}{p}\right)=\exp(\delta_1)=\varepsilon$. Finally as $y^p=\exp(p\cdot\log(y))>0$ it follows that $y^p\in B_{||}(0,\varepsilon)$ which proves that $[.]^p$ is continuous at 0 □

Now that we have extended the concept of powers we can prove Young' inequality. First we need a little lemma.

Lemma 18.41. *Let* $p\in]1,\infty[$ *and* $f_p\colon]0,\infty[\rightarrow\mathbb{R}$ *defined by* $f_p(x)=(1-x^{1-p})+(p-1)\cdot(1-x)$ *then* $x=1\Leftrightarrow f_p(x)=0$

Proof.

$\Rightarrow$. If $x=1$ then $f_p(1)=(1-1^{1-p})+(p-1)\cdot(1-1) \underset{18.39\ (2)}{=} (1-1)+(p-1)\cdot(1-1)=0$

⇐. Let $x \in]0, \infty[$ then $f_p'(x) = -(1-p) \cdot x^{-p} - (p-1) = (p-1) \cdot (x^{-p} - 1)$ and $f_p''(x) = (p-1) \cdot (-p) \cdot x^{-p} \underset{0<p-1, -p<0, 0<x^{-p}}{\Rightarrow} f_p''(x) < 0$ proving by 14.90 that f' is strictly decreasing on $]0, \infty[$

$$f_p'(1) = (p-1) \cdot (1^{-p} - 1) = (p-1) \cdot (1-1) = 0 \tag{18.20}$$

and

$$f' \text{ is strctly decreasing} \tag{18.21}$$

Assume now that $\exists x \in]0, \infty[$ such that $x \neq 1$ and $f_p(x) = 0$ we have either

$x < 1$. then as $f_p(x) = 0 = f_p(1)$ we have by the theorem of Rolle (see 14.88) that $\exists z \in]x, 1[$ such that $f_p'(z) = 0$. As f' is strictly increasing we must have that $f_p'(z) < f_p'(1) = 0$ giving $0 < 0$ a contradiction.

$1 < x$. hen as $f_p(1) = 0 = f_p(0)$ we have by the theorem of Rolle (see 14.88) that $\exists z \in]1, z[$ such that $f_p'(z) = 0$. As f' is strictly increasing we must have that $0 = f_p'(1) < f_p'(z)$ giving $0 < 0$ a contradiction.

As we have in both cases a contradiction we must have that $\forall x \in]0, \infty[$ with $x \neq 1$ that $f_p(x) \neq 0$. So if $f_p(x) = 0$ we must have $x = 1$ □

Theorem 18.42. (Young's inequality) *Let $p, q \in]0, \infty[$ such that $\frac{1}{p} + \frac{1}{q} = 1$ and $a, b \in]0, \infty[$ then we have that $a \cdot b \leqslant \frac{a^p}{p} + \frac{b^q}{q}$. Further if $p \in]0, \infty[$ then we have that $a \cdot b = \frac{a^p}{p} + \frac{b^q}{q} \Leftrightarrow b = a^{p-1} \Leftrightarrow a^p = b^q$*

Proof.

$$\begin{aligned}
a \cdot b &= \exp(\log(a \cdot b)) \\
&\underset{18.34}{=} \exp(\log(a) + \log(b)) \\
&= \exp\left(\frac{1}{p} \cdot p \cdot \log(a) + \frac{1}{q} \cdot q \cdot \log(b)\right) \\
&\underset{18.39}{=} \exp\left(\frac{1}{p} \cdot \log(a^p) + \frac{1}{q} \cdot \log(b^q)\right) \\
&\leqslant_{\text{exp is convex (see 18.28) and 14.93}} \frac{1}{p} \cdot \exp(\log(a^p)) + \frac{1}{q} \cdot \exp(\log(b^q)) \\
&= \frac{1}{p} \cdot a^p + \frac{1}{q} \cdot b^q \\
&= \frac{a^p}{p} + \frac{b^q}{q}
\end{aligned}$$

proving that

$$a \cdot b \leqslant \frac{a^p}{p} + \frac{b^q}{q} \tag{18.22}$$

Further we have that

$$\begin{aligned}\frac{1}{p}+\frac{1}{q}=1 &\Rightarrow \frac{p}{p}+\frac{p}{q}=p\\ &\Rightarrow 1+\frac{p}{q}=p\\ &\Rightarrow \frac{p}{q}=(p-1)\\ &\Rightarrow \frac{q}{p}+\frac{q}{q}=q\\ &\Rightarrow 1+\frac{q}{p}=q\\ &\Rightarrow \frac{q}{p}=(q-1)\end{aligned}$$

so we have by the hypothese that

$$\frac{p}{q}=(p-1)\wedge\frac{q}{p}=(q-1)\,\text{and thus}\,1=(p-1)\cdot(q-1) \tag{18.23}$$

Next we assume that additionally $p\in]1,\infty[$. First we have

$$\begin{aligned}a^p=b^q &\underset{0<q}{\Leftrightarrow} (a^p)^{\frac{1}{q}}=(b^q)^{\frac{1}{q}}\\ &\underset{18.39\ (5)}{\Leftrightarrow} a^{\frac{p}{q}}=b^{\frac{q}{q}}\\ &\Leftrightarrow a^{\frac{p}{q}}=b^1\\ &\Leftrightarrow a^{\frac{p}{q}}=b\\ &\underset{18.23}{\Leftrightarrow} a^{p-1}=b\end{aligned}$$

and

$$\begin{aligned}a^p=b^q &\underset{0<p}{\Leftrightarrow} (a^p)^{\frac{1}{p}}=(b^q)^{\frac{1}{p}}\\ &\underset{18.39\ (5)}{\Leftrightarrow} a^{\frac{p}{p}}=b^{\frac{q}{p}}\\ &\Leftrightarrow a^1=b^{\frac{q}{p}}\\ &\Leftrightarrow a=b^{\frac{q}{p}}\\ &\underset{18.23}{\Leftrightarrow} a=b^{q-1}\end{aligned}$$

so that we have

$$a^p=a^q\Leftrightarrow a^{p-1}=b\Leftrightarrow a=b^{\frac{q}{p}}\Leftrightarrow b=a^{\frac{p}{q}}\Leftrightarrow a=b^{q-1} \tag{18.24}$$

⇐. If $a=b^{p-1}$ then we have

$$\begin{aligned}\frac{a^p}{p}+\frac{b^q}{q} &= \frac{a^p\cdot q+b^q\cdot p}{p\cdot q}\\ &\underset{18.24}{=} \frac{q\cdot\left(b^{\frac{q}{p}}\right)^p+b^q\cdot p}{p\cdot q}\\ &= \frac{q\cdot b^{\frac{q}{p}\cdot p}+p\cdot b^q}{p\cdot q}\end{aligned}$$

$$
\begin{aligned}
&= \frac{q\cdot b^q + p\cdot b^q}{p\cdot q}\\
&= \frac{q+p}{p\cdot q}\cdot b^q\\
&= \left(\frac{1}{p}+\frac{1}{q}\right)\cdot b^q\\
&= b^q\\
&= b\cdot b^{q-1}\\
&\underset{18.24}{=} a\cdot b
\end{aligned}
$$

proving that $\frac{a^p}{p}+\frac{b^q}{q}=a\cdot b$

⇒. Suppose that $a\cdot b=\frac{a^p}{p}+\frac{b^q}{q}$ then we have

$$
\begin{aligned}
a\cdot b &= \frac{q\cdot a^p + p\cdot b^q}{p\cdot q}\\
&\underset{\frac{1}{p}+\frac{1}{q}=1\Rightarrow\frac{p+q}{p\cdot q}=1\Rightarrow\frac{1}{p\cdot q}=\frac{1}{p+q}}{=} \frac{q\cdot a^p + p\cdot b^q}{p+q}
\end{aligned}
$$

so that

$$
\begin{aligned}
p+q=\frac{q\cdot a^p+p\cdot b^q}{a\cdot b} &\Rightarrow p+q=q\cdot\frac{a^{p-1}}{b}+p\cdot\frac{b^{q-1}}{a}\\
&\Rightarrow p\cdot\left(1-\frac{b^{q-1}}{a}\right)=q\cdot\left(\frac{a^{p-1}}{b}-1\right)\\
&\Rightarrow \left(\frac{a^{p-1}}{b}-1\right)=\frac{p}{q}\cdot\left(1-\frac{b^{q-1}}{a}\right)\\
&\underset{18.23}{\Rightarrow} \left(\frac{a^{p-1}}{b}-1\right)=(p-1)\cdot\left(1-\frac{b^{q-1}}{a}\right)\\
&\underset{18.39(6)}{\Rightarrow} \left(\frac{b^{-1}}{a^{-(p-1)}}-1\right)=(p-1)\cdot\left(1-\frac{b^{q-1}}{a}\right)\\
&\underset{18.23}{\Rightarrow} \left(\frac{b^{-(p-1)\cdot(q-1)}}{a^{-(p-1)}}-1\right)=(p-1)\cdot\left(1-\frac{b^{q-1}}{a}\right)\\
&\Rightarrow \left(\frac{b^{(1-p)\cdot(q-1)}}{a^{(1-p)}}-1\right)=(p-1)\cdot\left(1-\frac{b^{q-1}}{a}\right)\\
&\Rightarrow \left(\left(\frac{b^{q-1}}{a}\right)^{1-p}-1\right)=(p-1)\cdot\left(1-\frac{b^{q-1}}{a}\right)\\
&\Rightarrow 0=\left(1-\left(\frac{b^{q-1}}{a}\right)^{1-p}\right)+(p-1)\cdot\left(1-\frac{b^{q-1}}{a}\right)
\end{aligned}
$$

so if we define $f_p:]0,\infty[\rightarrow\mathbb{R}$ by $f_p(x)=(1-x^{1-p})+(p-1)\cdot(1-x)$ we have by the above that $f_p\left(\frac{b^{q-1}}{a}\right)=0$. Using the previous lemma (see18.41) we conclude then that $\frac{b^{q-1}}{a}=1$ or that $a=b^{q-1}$ or using 18.23 that

$$b=a^{p-1}$$

□

Next we examine the relation between the exponential map and goniometric functions, for that we must consider the behavior of exp on the space of complex numbers.

Theorem 18.43. $\exp: \mathbb{C} \rightarrow \mathbb{C}$ *has the following properties*

1. $\forall z \in \mathbb{C}$ *we have that* $\overline{\exp(z)} = \exp(z)$
2. $\forall z \in \mathbb{C}$ *we have* $|\exp(z)|^2 = \exp(\mathrm{Re}(z))$
3. $\forall z = x + i \cdot y \in \mathbb{C}$ *we have that* $|\exp(z)| = \exp(x)$ *and* $|\exp(i \cdot y)| = \exp(0) = 1$

Proof.

1.

$$\begin{aligned}
\exp(\bar{z}) &= \sum_{i=0}^{\infty} \frac{\bar{z}^i}{i!} \\
&= \lim_{n \rightarrow \infty} \left(\sum_{i=0}^{n} \frac{\bar{z}^i}{i!} \right) \\
&= \lim_{n \rightarrow \infty} \left(\sum_{i=0}^{n} \frac{\overline{z^i}}{i!} \right) \\
&= \lim_{n \rightarrow \infty} \overline{\sum_{i=0}^{\infty} \frac{z^i}{i!}} \\
&\underset{12.322}{=} \overline{\lim_{n \rightarrow \infty} \left(\sum_{i=0}^{\infty} \frac{z^i}{i!} \right)} \\
&= \exp(z)
\end{aligned}$$

2.

$$\begin{aligned}
|\exp(z)|^2 &= \exp(z) \cdot \overline{\exp(z)} \\
&\underset{(1)}{=} \exp(z) \cdot \exp(\bar{z}) \\
&= \exp(z + \bar{z}) \\
&= \exp(2 \cdot \mathrm{Re}(z)) \\
&= \exp(\mathrm{Re}(z))^2
\end{aligned}$$

Now as $|\exp(z)|, \exp(\mathrm{Re}(z)) \in \mathbb{R}_+$ we have by 9.70 that $|\exp(z)| = \exp(\mathrm{Re}(z))$

3. If $x = x + i \cdot y$ then $|\exp(z)| \underset{(2)}{=} \exp(\mathrm{Re}(z)) = \exp(x)$. Further $|\exp(i \cdot y)| = \exp(\mathrm{Re}(i \cdot y)) = \exp(0) = 1$ □

We can define the trigoniometric functions in terms of the exponential function

Definition 18.44. (trigoniometric functions)

1. $\cos: \mathbb{C} \rightarrow \mathbb{C}$ *is defined by* $\cos = \frac{\exp(i \cdot z) + \exp(-i \cdot z)}{2}$
2. $\sin: \mathbb{C} \rightarrow \mathbb{C}$ *is defined by* $\sin = \frac{\exp(i \cdot z) - \exp(-i \cdot z)}{2 \cdot i}$

Theorem 18.45. *We have the following properties for the trigoniometric functions*

1. $\cos: \mathbb{C} \rightarrow \mathbb{C}$ *and* $\sin: \mathbb{C} \rightarrow \mathbb{C}$ *are continuous on* $\mathbb{C}$
2. $\forall z \in \mathbb{C}$ *we have that* $\cos'(z) = -\sin(z)$ *and* $\sin'(z) = \cos(z)$
3. $\cos(0) = 1$
4. $\sin(0) = 0$
5. $\forall z \in \mathbb{C}$ *we have*
   a. $\cos(-z) = \cos(-z)$
   b. $\sin(-z) = -\sin(z)$
6. $\forall z \in \mathbb{C}$ *we have* $\cos(z)^2 + \sin(z)^2 = 1$
7. $\forall x, y \in \mathbb{C}$ *we have*
   a. $\cos(x + y) = \cos(x) \cdot \cos(y) - \sin(x) \cdot \sin(y)$
   b. $\sin(x + y) = \sin(x) \cdot \cos(y) + \cos(x) \cdot \sin(y)$
8. $\forall x \in \mathbb{R}$ *we have*
   a. $\cos(x) \in \mathbb{R}$
   b. $\sin(x) \in \mathbb{R}$
9. $\forall x \in \mathbb{C}$ *we have* $e^{i \cdot x} = \cos(x) + i \cdot \sin(x)$ *(which by (8) means that for* $x \in \mathbb{R}$ *we have* $\mathrm{Re}(i \cdot x) = \cos(x)$ *and* $\mathrm{Im}(e^{i \cdot x}) = \sin(x)$
10. $\forall x \in \mathbb{R}$ *we have* $|e^{i \cdot x}| = 1$

Proof.

1. This follows as the product of a scalar with a continuous function is continuous and the sum of continuous functions is continuous.

2. Let $z \in \mathbb{C}$ then

$$\begin{aligned}
\cos'(z) &= \frac{1}{2} \cdot (i \cdot \exp(i \cdot z) + (-i) \cdot \exp(-i \cdot z)) \\
&= \frac{i \cdot i \cdot \exp(i \cdot z) + i \cdot (-i) \cdot \exp(-i \cdot z)}{2 \cdot i} \\
&= \frac{-\exp(i \cdot z) + \exp(-i \cdot z)}{2 \cdot i} \\
&= -\sin(z)
\end{aligned}$$

and

$$\begin{aligned}
\sin'(z) &= \frac{1}{2_{\cdot i}} \cdot (i \cdot \exp(i \cdot z) - (-i) \cdot \exp(-i \cdot z)) \\
&= \frac{\exp(i \cdot z) + \exp(i \cdot z)}{2} \\
&= \cos(z)
\end{aligned}$$

3.

$$\begin{aligned}
\cos(0) &= \frac{\exp(i \cdot 0) + \exp(-i \cdot 0)}{2} \\
&= \frac{1+1}{2} \\
&= 1
\end{aligned}$$

4.

$$\begin{aligned}
\sin(0) &= \frac{\exp(i \cdot 0) - \exp(-i \cdot 0)}{2 \cdot i} \\
&= \frac{1-1}{2 \cdot i} \\
&= 0
\end{aligned}$$

5.

a.

$$\begin{aligned}
\cos(-z) &= \frac{\exp(i \cdot (-z)) + \exp(-i \cdot (-z))}{2} \\
&= \frac{\exp(i * z) + \exp(-i \cdot z)}{2} \\
&= \cos(z)
\end{aligned}$$

b.

$$\begin{aligned}
\sin(-z) &= \frac{\exp(i \cdot (-z)) - \exp(-i \cdot (-z))}{2 \cdot i} \\
&= \frac{\exp(-i \cdot z) - \exp(i \cdot z)}{2 \cdot i} \\
&= -\sin(z)
\end{aligned}$$

6.

$$\begin{aligned}
\cos(z)^2+\sin(z)^2 &= \frac{(\exp(i\cdot z)+\exp(-i\cdot z))^2}{4}+\frac{(\exp(i\cdot z)-\exp(-i\cdot z))^2}{-4}\\
&= \frac{1}{4}\cdot((\exp(i))^2+(\exp(-i\cdot z))^2+2\cdot\exp(i\cdot z)\cdot\exp(-i\cdot z)-\\
&\quad (\exp(i\cdot z))^2-(\exp(-i\cdot z))^2+2\cdot\exp(i\cdot z)\cdot\exp(-i\cdot z))\\
&= \frac{1}{4}\cdot(4\cdot\exp(i\cdot z)\cdot\exp(-i\cdot z))\\
&= \exp(i\cdot z)\cdot\exp(-i\cdot z)\\
&= \exp(i\cdot z-i\cdot z)\\
&= \exp(0)\\
&= 1
\end{aligned}$$

7.

a.

$$\begin{aligned}
\cos(x)\cdot\cos(y)-\sin(x)\cdot\sin(y) &= \frac{\exp(i\cdot x)+\exp(-i\cdot x)}{2}\cdot\\
&\quad \frac{\exp(i\cdot y)+\exp(-i\cdot y)}{2}-\\
&\quad \frac{\exp(i\cdot x)-\exp(-i\cdot x)}{2\cdot i}\cdot\\
&\quad \frac{\exp(i\cdot y)-\exp(-i\cdot y)}{2\cdot i}\\
&= \frac{1}{4}\cdot(\exp(i\cdot x)\cdot\exp(i\cdot y)+\exp(i\cdot x)\cdot\exp(-i\cdot y)+\exp(-i\cdot x)\cdot\exp(i\cdot y)+\exp(-i\cdot x)\cdot\exp(-i\cdot y))+\frac{1}{4}\cdot\\
&\quad (\exp(i\cdot x)\cdot\exp(i\cdot y)-\exp(i\cdot x)\cdot\exp(-i\cdot y)-\exp(-i\cdot x)\cdot\exp(i\cdot y)+\exp(-i\cdot x)\cdot\exp(-i\cdot y))\\
&= \frac{1}{4}\cdot(\exp(i\cdot x)\cdot\exp(i\cdot y)+\exp(i\cdot x)\cdot\exp(-i\cdot y)+\exp(-i\cdot x)\cdot\exp(i\cdot y)+\exp(-i\cdot x)\cdot\exp(-i\cdot y)+\\
&\quad \exp(i\cdot x)\cdot\exp(i\cdot y)-\exp(i\cdot x)\cdot\exp(-i\cdot y)-\exp(-i\cdot x)\cdot\exp(i\cdot y)+\exp(-i\cdot x)\cdot\exp(-i\cdot y))\\
&= \frac{1}{4}\cdot(2\cdot\exp(i\cdot x)\cdot\exp(i\cdot y)+2\cdot\exp(-i\cdot x)\cdot\exp(-i\cdot y))\\
&= \frac{1}{2}\cdot(\exp(i\cdot x+i\cdot y)+\exp(-i\cdot x-i\cdot y))
\end{aligned}$$

$$\begin{aligned}
&= \frac{\exp(i\cdot(x+y))+\exp(-i\cdot(x+y))}{2} \\
&= \cos(x+y)
\end{aligned}$$

b.

$$\begin{aligned}
\sin(x)\cdot\cos(y)+\cos(x)\cdot\sin(y) &= \frac{\exp(i\cdot x)-\exp(-i\cdot x)}{2\cdot i}\cdot \frac{\exp(i\cdot y)+\exp(-i\cdot y)}{2} + \frac{\exp(i\cdot x)+\exp(-i\cdot x)}{2}\cdot \frac{\exp(i\cdot y)-\exp(-i\cdot y)}{2\cdot i} \\
&= \frac{1}{4\cdot i}\cdot(\exp(i\cdot x)\cdot\exp(i\cdot y)+\exp(i\cdot x)\cdot\exp(-i\cdot y)-\exp(-i\cdot x)\cdot\exp(i\cdot y)-\exp(-i\cdot x)\cdot\exp(-i\cdot y)+\exp(i\cdot x)\cdot\exp(i\cdot y)-\exp(i\cdot x)\cdot\exp(-i\cdot y)+\exp(-i\cdot x)\cdot\exp(i\cdot y)-\exp(-i\cdot x)\cdot\exp(-i\cdot y)) \\
&= \frac{1}{4\cdot i}\cdot(2\cdot\exp(i\cdot x)\cdot\exp(i\cdot y)-2\cdot\exp(-i\cdot x)\cdot\exp(-i\cdot y)) \\
&= \frac{\exp(i\cdot x+i\cdot y)-\exp(-i\cdot x-i\cdot y)}{2\cdot i} \\
&= \frac{\exp(i\cdot(x+y))-\exp(-i\cdot(x+y))}{2\cdot i} \\
&= \sin(x+y)
\end{aligned}$$

8. Let $x\in\mathbb{R}$ then

a.

$$\begin{aligned}
\cos(x) &= \frac{\exp(i\cdot x)+\exp(-i\cdot x)}{2} \\
&= \frac{\exp(i\cdot x)+\exp(\overline{i\cdot x})}{2} \\
&\underset{18.43\ (1)}{=} \frac{\exp(i\cdot x)+\overline{\exp(i\cdot x)}}{2} \\
&= \frac{2\cdot\mathrm{Re}(\exp(i\cdot x))}{2} \\
&= \mathrm{Re}(\exp(i\cdot x)) \\
&\in \mathbb{R}
\end{aligned}$$

b.

$$\begin{aligned}
\sin(x) &= \frac{\exp(i\cdot x)-\exp(-i\cdot x)}{2\cdot i} \\
&= \frac{\exp(i\cdot x)-\exp(\overline{i\cdot x})}{2\cdot i}
\end{aligned}$$

$$\begin{aligned}
&= \frac{\exp(i\cdot x) - \exp(\overline{i\cdot x})}{2\cdot i}\\
&= \frac{2\cdot i\cdot \mathrm{Img}(\exp(i\cdot x))}{2\cdot i}\\
&= \mathrm{Img}(\exp(i\cdot x))\\
&\in \mathbb{R}
\end{aligned}$$

9. Let $x\in\mathbb{R}$ then

$$\begin{aligned}
\cos(x) + i\cdot\sin(x) &= \frac{\exp(i\cdot x)+\exp(-i\cdot x)}{2} + i\cdot\frac{\exp(i\cdot x)-\exp(-i\cdot x)}{2\cdot i}\\
&= \frac{1}{2}\cdot(\exp(i\cdot x)+\exp(-i\cdot x)+\exp(i\cdot x)-\exp(-i\cdot x))\\
&= \exp(i\cdot x)
\end{aligned}$$

10. Let $x\in\mathbb{R}$ then

$$\begin{aligned}
|e^{i\cdot x}| &\underset{(9)}{=} |\cos(x)+i\cdot\sin(x)|\\
&\underset{\cos(x),\sin(x)\in\mathbb{R}}{=} \cos^2(x)+\sin^2(x)\\
&\underset{(6)}{=} 1
\end{aligned}$$

□

The next theorem shows that $\cos:\mathbb{R}\rightarrow\mathbb{R}$ has a zero, this will allows us to define π and show thqat sin and cos are periodically.

Theorem 18.46. $\exists x_0\in[0,\infty[$ *such that* $\cos(x)=0$

Proof. We proof this by contradiction so let assume that $\forall x\in[0,\infty[$ we have that $\cos(x)\neq 0$. If now there exists a $x\in[0,\infty[$ such that $\cos(x)\leqslant 0<1 \underset{18.45\ (3)}{=} \cos(0)\Rightarrow 0\in[\cos(x),1]$ we have by the continuity of exp and the intermediate value theorem (see 12.444) that there exists a $y\in[x,1]$ such that $\cos(y)=0$ contradicting the assumption that $\cos(x)\neq 0$. So we conclude that $\forall x\in[0,\infty[$ we have $0<\cos(x)$, as $\sin'(x)=\cos(x)$ we conclude that $\forall x\in[0,\infty[\ \sin'(x)>0$ hence by 14.90 we have that

$$\sin:[0,\infty[\rightarrow[0,\infty[\text{ is strictly increasing} \tag{18.25}$$

As $\sin(0) \underset{18.44\ (4)}{=} 0$ we conclude that

$$\forall x\in]0,\infty[\text{ that } 0<\sin(x) \tag{18.26}$$

Further we have by 18.45 (6) that $\forall x\in\mathbb{R}$ that $\cos^2(x)+\sin^2(x)=1$ so that $|\cos(x)|^2=\cos^2(x)=1-\sin^2(x)\leqslant 1$ so as $0\leqslant|\cos(x)|$ and $\sqrt{}$ is increasing (see 9.71) we have that $|\cos(x)|=\sqrt{|\cos(x)|^2}\leqslant\sqrt{1}=1$ so that

$$\forall x\in\mathbb{R} \text{ we have } \cos(x), -\cos(x)\leqslant 1 \tag{18.27}$$

We have then for $0<x<y$ that if we take $C_{\sin(x)}:[0,x]\rightarrow\mathbb{R}$ defined by $C_{\sin(x)}(t)=\sin(x)$ so that $\forall t\in[x,y]$ we have $C_{\sin(x)}(t)\underset{18.25}{=}\sin(x)<\sin(t)$

$$\begin{array}{lcl}\sin(x)\cdot(y-x) & \underset{12.433}{=} & \int_x^y C_{\sin(x)}\\ & \leqslant_{C_{\sin(x)}(t)\leqslant\sin(t)\text{ and }12.429} & \int_x^y \sin\\ & \underset{\cos'=-\sin}{=} & -\int_x^y \cos'\\ & \underset{\text{fundamental theorem of calculus (14.98)}}{=} & -(\cos(y)-\cos(x))\\ & = & \cos(x)-\cos(y)\\ & \leqslant_{18.27} & 2\end{array}$$

proving that

$$\forall x,y\in\mathbb{R}\text{ with }0<x<y\text{ we have }(y-x)\leqslant\frac{2}{\sin(x)} \tag{18.28}$$

Take now $0<x$ and take $y=\frac{2}{\sin(x)}+x+1$ then as $0<\sin(x)$ we have that $0<x<y$ and $y-x=\frac{2}{\sin(x)}+1>\frac{2}{\sin(x)}\geqslant_{18.28}(y-x)$ a contradiction. So our assumption is wrong and ther must exists a $x\in[0,\infty[$ such that $\cos(x)=0$ $\square$

Corollary 18.47. $\{x\in[0,.\infty[|\cos(x)=0\}=\cos^{-1}(\{0\})\bigcap[0,\infty[$ *has a minimum*

Proof. Using the previous theorem we we have that $\cos^{-1}(\{0\})\bigcap[0,\infty[\neq\emptyset$ so that by the conditionally completeness of the real numbers (see 9.43) $x_0=\cos^{-1}(\{0\})$ exists. Further as $\{0\}$ is closed (see 12.220), $[0,\infty[$ is closed (see 12.72) and the continuity of cos it follows that $\cos^{-1}(\{0\})\bigcap[0,\infty[$ is closed. Assume now that $x_0\notin\cos^{-1}(\{0\})\bigcap[0,\infty[$ then $x_0\in\mathbb{R}\setminus(\cos^{-1}(\{0\})\bigcap[0,\infty[)$ a open set so that $\exists\delta>0$ such that $]x_0-\delta,x_0+\delta[=B_{||}(x_0,\delta)\subseteq\mathbb{R}\setminus(\cos^{-1}(\{0\})\bigcap[0,\infty[)$ proving that $]x_0-\delta,x_0+\delta[\bigcap\cos^{-1}(\{0\})\bigcap[0,\infty[=\emptyset$. As $x_0<x_0+\delta$ there exists by the definition of a infinum a $y\in\cos^{-1}(\{0\})\bigcap[0,\infty[$ such that $x_0\leqslant y<x_0+\delta$ proving that $(\cos^{-1}(\{0\})\bigcap[0,\infty[)\bigcap]x_0-\delta,x_0+\delta[\neq\emptyset$ a contradiction. So we must have that $x_0\in\cos^{-1}(\{0\})\bigcap[0,\infty[$ which as $\forall y\in\cos^{-1}(\{0\})\bigcap[0,\infty[$ we have $x_0\leqslant y$ means that $x_0=\min(\cos^{-1}(\{0\})\bigcap[0,\infty[)$ $\square$

The above corollary allows us to define π for the trigoniometric functions.

Definition 18.48. **(π)** $\pi=2\cdot\min(\cos^{-1}(\{0\})\bigcap[0,\infty[)$ *so that* $\frac{\pi}{2}=\min(\cos^{-1}(\{0\})\bigcap[0,\infty[)$ *hence* $\cos\left(\frac{\pi}{2}\right)=0$ *and* $\forall x\in\left[0,\frac{\pi}{2}\right[$ *we have* $\cos\left(\frac{\pi}{2}\right)\neq0$. *Note that as* $\cos(0)=1$ *we have that* $0<\frac{\pi}{2}$.

Next we prove that π and the trigoniometric functions have the usual properties

Theorem 18.49. We have

1. $\forall x\in\left[0,\frac{\pi}{2}\right[$ *we have* $0<\cos(x)$ *(hence as* $\cos\left(\frac{\pi}{2}\right)=0$ *we have* $\forall x\in\left[0,\frac{\pi}{2}\right]$ *that* $0\leqslant\cos(x)$*)*

2. *$\forall x \in \left]0, \frac{\pi}{2}\right]$ we have $0 < \sin(x)$ (hence as $\sin(0) = 0$ we have $\forall x \in \left[0, \frac{\pi}{2}\right]$ that $0 \leqslant \sin(x)$)*

Proof.

1. Assume that $\exists y \in \left[0, \frac{\pi}{2}\right[$ such that $\cos(y) \leqslant 0$ then $\cos(y) < 1 \underset{18.45\ (3)}{=} \cos(0)$ so using the intermediate value theorem (see 12.444) there exists a $z \in [0, y]$ such that $\cos(z) = 0$. Hence as $y < \frac{\pi}{2}$ we have found a $z \in \left[0, \frac{\pi}{2}\right[$ such that $\cos(z) = 0$ contradicting $\frac{\pi}{2} = \min(\cos^{-1}(\{0\}) \bigcap [0, \infty[)$. So we must have that $\forall y \in \left[0, \frac{\pi}{2}\right[$ that $0 < \cos(y)$

2. As $\forall y \in \left[0, \frac{\pi}{2}\right[$ we have $\sin'(x) \underset{18.45\ (2)}{=} \cos(x) > 0$ (by (1)) we have by 14.90 that sin is strictly increasing on $\left[0, \frac{\pi}{2}\right[$ so that

$$\forall x \in \left]0, \frac{\pi}{2}\right[\text{ we have } 0 < \sin(x) \tag{18.29}$$

More special we have that for $0 < \frac{\pi}{4} < \frac{\pi}{2}$ that $0 < \sin\left(\frac{\pi}{4}\right)$, then as $\forall x \in \left[\frac{\pi}{4}, \frac{\pi}{2}\right]$ we have by (2) that $\sin'(x) = \cos(x) \geqslant 0$ that by 14.90 $\sin(x)$ is increasing on

□

Theorem 18.50. *We have the following properties for π and the trigoniometric functions*

1. $\cos(0) = 1$
2. $\cos\left(\frac{\pi}{2}\right) = 0$
3. $\sin(0) = 0$
4. $\sin\left(\frac{\pi}{2}\right) = 1$
5. $\cos(\pi) = -1$
6. $\sin(\pi) = 0$
7. $\cos(2 \cdot \pi) = 1$
8. $\sin(2 \cdot \pi) = 0$
9. $\forall x \in \mathbb{R}$ *we have*
   a. $\sin(x + 2 \cdot \pi) = \sin(x)$
   b. $\cos(x + 2 \cdot p) = \cos(x)$
10. $\forall x \in \mathbb{R}$ *we have* $\exp(i \cdot (x + 2 \cdot \pi)) = \exp(i \cdot x)$ *(or using the power notation* $e^{i \cdot (x + 2 \cdot \pi)} = e^{i \cdot x}$*)*
11. $\exp(i \cdot \pi) = -1$ *(or using the power notation* $e^{i \cdot \pi} = -1$*) (the most beautifull equation in the world because it shows how three funcamental constants* (π, e *and* 1) *are related)*

Proof.

1. This was already proved in 18.45 (3)

2. This follows from the definition of $\frac{\pi}{2}$
3. This was already proved in 18.45 (4)
4. Using 18.45 (6) we have that $\cos^2\left(\frac{\pi}{2}\right)+\sin^2\left(\frac{\pi}{2}\right)=1 \underset{(2)}{\Rightarrow} \sin^2\left(\frac{\pi}{2}\right)=1$ so that $\sin\left(\frac{\pi}{2}\right)=1 \vee \sin\left(\frac{\pi}{2}\right)=-1$, as $\sin\left(\frac{\pi}{2}\right)>1$ (see 18.49) it follows that $\sin\left(\frac{\pi}{2}\right)=1$
5.

$$\begin{array}{rcl} \cos(\pi) & = & \cos\left(\frac{\pi}{2}+\frac{\pi}{2}\right) \\ & \underset{18.45\ (7)}{=} & \cos\left(\frac{\pi}{2}\right)\cdot\cos\left(\frac{\pi}{2}\right)-\sin\left(\frac{\pi}{2}\right)\cdot\sin\left(\frac{\pi}{2}\right) \\ & \underset{(2)\,\text{and}\,(4)}{=} & 0\cdot 0-1\cdot 1 \\ & = & -1 \end{array}$$

6.

$$\begin{array}{rcl} \sin(\pi) & = & \sin\left(\frac{\pi}{2}+\frac{\pi}{2}\right) \\ & \underset{18.45\ (7)}{=} & \sin\left(\frac{\pi}{2}\right)\cdot\cos\left(\frac{\pi}{2}\right)+\cos\left(\frac{\pi}{2}\right)\cdot\sin\left(\frac{\pi}{2}\right) \\ & \underset{(2)\,\text{and}\,(4)}{=} & 1\cdot 0-0\cdot 1 \\ & = & 0 \end{array}$$

7.

$$\begin{array}{rcl} \cos(2\cdot\pi) & = & \cos(\pi+\pi) \\ & = & \cos(\pi)\cdot\cos(\pi)-\sin(\pi)\cdot\sin(\pi) \\ & \underset{(5)\,\text{and}\,(6)}{=} & (-1)\cdot(-1)-0\cdot 0 \\ & = & 1 \end{array}$$

8.

$$\begin{array}{rcl} \sin(2\cdot\pi) & = & \sin(\pi+\pi) \\ & \underset{18.45\ (7)}{=} & \sin(\pi)\cdot\cos(\pi)+\cos(\pi)\cdot\sin(\pi) \\ & \underset{(5)\,\text{and}\,(6)}{=} & 0\cdot(-1)+(-1)\cdot 0 \\ & = & 0 \end{array}$$

9. Let $x\in\mathbb{R}$ then

a.
$$\begin{array}{rcl} \cos(x+2\cdot\pi) & \underset{18.45\ (7)}{=} & \cos(x)\cdot\cos(2\cdot\pi)-\sin(x)\cdot\sin(2\cdot\pi) \\ & \underset{(7)\,\text{and}\,(8)}{=} & \cos(x)\cdot 1-\sin(x)\cdot 0 \\ & = & \cos(x) \end{array}$$

b.
$$\begin{array}{rcl} \sin(x+2\cdot\pi) & \underset{18.45\ (7)}{=} & \sin(x)\cdot\cos(2\cdot\pi)+\cos(x)\cdot\sin(2\cdot\pi) \\ & \underset{(7)\,\text{and}\,(8)}{=} & \sin(x)\cdot 1+\cos(x)\cdot 0 \\ & = & \sin(x) \end{array}$$

10. Let $x \in \mathbb{R}$ then

$$\begin{array}{rcl} \exp(i \cdot (x + 2 \cdot \pi)) & = & \cos(x + 2 \cdot \pi) + i \cdot \sin(x + 2 \cdot \pi) \\ & \underset{(9)}{=} & \cos(x) + i \cdot \sin(x) \\ & = & \exp(i \cdot x) \end{array}$$

11.

$$\begin{array}{rcl} \exp(i \cdot \pi) & = & \cos(\pi) + i \cdot \sin(\pi) \\ & \underset{(5)\,(6)}{=} & -1 + i \cdot 0 \\ & = & -1 \end{array}$$

$\square$

Chapter 18

Measure Theory

Convention 18.1. In this chapter we use the following conventions

1. If $A \subseteq \mathbb{R}$ then $\sup(A) < \infty$ means that $\sup(A)$ exists in $\mathbb{R}$ [see Convention 17.14]
2. If $A \subseteq \mathbb{R}$ then $-\infty < \inf(A)$ means that $\inf(A)$ exists in $\mathbb{R}$ [see Convention 17.14]
3. Let $k \in \mathbb{N}$ and $\{x_i\}_{i \in \{k, \ldots, \infty\}} \subseteq \mathbb{R}$ then $\lim_{n \to \infty} x_n \in \mathbb{R}$ means that $\lim_{n \to \infty} x_n$ exists in $\mathbb{R}$ [see also 17.76]
4. Let $\{x_i\}_{i \in \mathbb{N}} \subseteq \mathbb{R}$ then $\liminf_{n \to \infty} x_n \in \mathbb{R}$ means that $\sup(\{\inf(\{x_i | i \in \{n, \ldots, \infty\}\}) | n \in \mathbb{N}\}) \in \mathbb{R}$
5. Let $\{x_i\}_{i \in \mathbb{N}} \subseteq \mathbb{R}$ then $\limsup_{n \to \infty} x_n \in \mathbb{R}$ means that $\inf(\{\sup(\{x_i | i \in \{n, \ldots, \infty\}\}) | n \in \mathbb{N}\}) \in \mathbb{R}$

Notation 18.2. *Given $a, b \in \bar{\mathbb{R}}$ we use the following notations*

1. $[a, b] = \{x \in \bar{\mathbb{R}} | a \leqslant x \leqslant b\}$
2. $[a, b[= \{x \in \bar{\mathbb{R}} | a \leqslant x < b\}$
3. $]a, b] = \{x \in \bar{\mathbb{R}} | a < x \leqslant b\}$
4. $]a, b[= \{x \in \bar{\mathbb{R}} | a < x < b\}$ *[note that this set is guaranteed to be a subset of $\mathbb{R}$]*

the special cases where $a = -\infty$ or $b = \infty$ are called segments.

18.1 Basic concepts of measure theory

18.1.1 Measurable spaces

Measure theory is involved with the concepts of length, area or volume of sets, actually a measure of a set contains is the unification of the terms 'length', 'area' and 'volume' of a set. To apply this term of measure in a meaningful way on sets and avoid paradoxes (like the Tarski paradox) we must restrict ourselves to certain classes of sets. This is the idea of σ-algebra's.

Definition 18.3. (σ-algebra) *Let X be a set, $\mathcal{A} \subseteq \mathcal{P}(X)$ [$\mathcal{P}(X)$ is the set of subsets of X (see 1.68)] then $\mathcal{A}$ is a a σ-algebra on X iff*

1. $\emptyset \in \mathcal{A}$
2. $\forall A \in \mathcal{A}$ *we have that* $X \setminus A \in \mathcal{A}$

3. *$\forall\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{A}$ [meaning that $\forall i\in\mathbb{N}$ we have $A_i\in\mathcal{A}$] we have that $\bigcup_{i\in\mathbb{N}}A_i\in\mathcal{A}$*

Notation 18.4. (Measurable Space) *Let X be a set and $\mathcal{A}$ a σ-algebra then the structure $\langle X,\mathcal{A}\rangle$ is called a **measurable space**.*

A trivial example of a measure space is the following, although it is not very useful

Example 18.5. Let X be a set then $\langle X,\mathcal{P}(X)\rangle$ is a measurable space

Proof.

1. $\emptyset\in\mathcal{P}(X)$
2. If $A\in\mathcal{P}(X)$ then $X\setminus A\subseteq X$ so that $X\setminus A\in\mathcal{P}(X)$
3. $\forall\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{P}(X)$ we have $A_i\subseteq X$ so that $\bigcup_{i\in\mathbb{N}}A_i\subseteq X$ proving that $\bigcup_{i\in\mathbb{N}}A_i\in\mathcal{P}(X)$ □

Actually the above definition can be rewritten using the more general concept of countable sets (see 5.30) as is seen in the following proposition.

Proposition 18.6. *Let X be a set, $\mathcal{A}\subseteq\mathcal{P}(X)$ then $\mathcal{A}\subseteq\mathcal{P}(X)$ is a σ-algebra iff*

1. *$\emptyset\in\mathcal{A}$*
2. *$\forall A\in\mathcal{A}$ we have $X\setminus A\in\mathcal{A}$*
3. *$\forall\{A_i\}_{i\in K}\subseteq\mathcal{A}$ where K is countable we have that $\bigcup_{i\in K}A_i\in\mathcal{A}$*

Proof.

⇒. Let $\mathcal{A}$ be a algebra then (1) and (2) in the proposition follows from the definition of a σ-algebra. So we must only prove (3), so let K be a countable set then we have the following possibilities to consider:

***K* is finite.** then either $K=\emptyset$ and then $\bigcup_{i\in K}A_i\underset{1.100}{=}\emptyset\in\mathcal{A}$ or there exists a $n\in\mathbb{N}$ and a bijection $\beta:\{1,\ldots,n\}\to K$ define then $\{B_i\}_{i\in\mathbb{N}}$ by $B_i=\begin{cases}A_{\beta(i)}\text{ if } i\in\{1,\ldots,n\}\\ \emptyset\text{ if } i\in\mathbb{N}\setminus\{1,\ldots,n\}\end{cases}$ then we have

$$\begin{aligned}
x\in\bigcup_{i\in\mathbb{N}}B_i &\Rightarrow \exists i\in\mathbb{N}\vDash x\in B_i\\
&\underset{i\in\mathbb{N}\setminus\{1,\ldots,n\}\Rightarrow x\in B_i=\emptyset\text{ a contradiction}}{\Rightarrow} i\in\{1,\ldots,n\}\\
&\Rightarrow x\in B_i=A_{\beta(i)}\subseteq_{\beta(i)\in K}\bigcup_{i\in K}A_i
\end{aligned}$$

$$\begin{aligned}
x\in\bigcup_{i\in K}A_i &\Rightarrow \exists i\in K\vDash x\in A_I\\
&\underset{\beta\text{ is surjective}}{\Rightarrow} \exists j\in\{1,\ldots,n\}\subseteq\mathbb{N}\vDash\beta(j)=i\\
&\Rightarrow x\in A_{\beta(j)}=B_j\subseteq\bigcup_{i\in\mathbb{N}}B_i
\end{aligned}$$

which proves that $\bigcup_{i \in K} A_i = \bigcup_{i \in \mathbb{N}} B_i \in \mathcal{A}$ [as $\mathcal{A}$ is a σ-algebra]

***K* is infinite countable.** then there exists a bijection $\beta: \mathbb{N} \to K$, define then $\{B_i\}_{i \in \mathbb{N}}$ by $B_i = A_{\beta(i)}$. Then

$$\begin{array}{rcl} x \in \bigcup_{i \in \mathbb{N}} B_i & \Rightarrow & \exists i \in \mathbb{N} \vDash x \in B_i = A_{\beta(i)} \subseteq \bigcup_{k \in K} B_k \\ x \in \bigcup_{k \in K} A_k & \Rightarrow & \exists k \in K \vDash x \in A_k \\ & \underset{\beta \text{ is bijection}}{\Rightarrow} & \exists i \in \mathbb{N} \vDash \beta(i) = k \\ & \Rightarrow & x \in A_k = A_{\beta(i)} = B_i \subseteq \bigcup_{i \in \mathbb{N}} B_i \end{array}$$

which proves that $\bigcup_{i \in K} A_i = \bigcup_{i \in \mathbb{N}} B_i \in \mathcal{A}$

So in all cases we have that $\bigcup_{i \in K} A_i \in \mathcal{A}$ proving that (3)

$\Leftarrow$. If $\mathcal{A}$ satisfies (1),(2) of the proposition then (1) (2) of the definition of a σ-algebra is satisfied and (3) is satisfied because $\mathbb{N}$ is countable. □

A trivial example of a σ-algebra is the set of all subsets of X although it is not a very useful one.

Example 18.7. Let X be a set then $\mathcal{P}(X)$ is a σ-algebra on X

Proof.

1. As $\emptyset \subseteq X$ we have $\emptyset \in \mathcal{P}(X)$
2. If $A \in \mathcal{P}(X)$ then $X \setminus A \subseteq X$ so that $X \setminus A \in \mathcal{P}(X)$
3. If $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{P}(X)$ then $\forall i \in \mathbb{N}$ we have $A_i \subseteq \mathcal{P}(X)$ so that $\bigcup_{i \in \mathbb{N}} A_i \subseteq X$ proving that $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{P}(X)$ □

Proposition 18.8. *Let $\langle X, \mathcal{A} \rangle$ be a measure space then we have*

1. *$\forall \{A_i\}_{i \in K} \subseteq \mathcal{A}$ with K countable we have $\bigcap_{i \in K} A_i \in \mathcal{A}$*
2. *$\forall A, B \in \mathcal{A}$ we have*
   a. *$A \bigcup B \in \mathcal{A}$*
   b. *$A \bigcap B \in \mathcal{A}$*
   c. *$A \setminus B \in \mathcal{A}$*
3. *$X \in \mathcal{A}$*

Proof.

1. Define $\{B_i\}_{i \in K}$ by $B_i = X \setminus A_i$ then using 18.6 (2) we have that $\{B_i\}_{i \in K} \subseteq \mathcal{A}$ hence by 18.6 (3) $\bigcup_{i \in K} B_i \in \mathcal{A}$ and thus

$$X \setminus \left(\bigcup_{i \in K} B_i \right) \in \mathcal{A} \tag{18.1}$$

Further we have

$$X \setminus \left(\bigcup_{i\in K} B_i\right) \underset{1.108}{=} \bigcap_{i\in K} (X\setminus B_i)$$
$$= \bigcap_{i\in K} (X\setminus(X\setminus A_i))$$
$$\underset{1.31}{=} \bigcap_{i\in K} A_i$$

so that $\bigcap_{i\in K} A_i \in \mathcal{A}$

2. Let $A, B \in \mathcal{A}$

   a. Define $\{A_i\}_{i\in\{1,2\}}$ by $A_1 = A \wedge A_2 = B$ then as $\{1,2\}$ is finite hence countable $A \bigcup B \underset{18.6\,(3)}{=} \bigcup_{i\in K} A_i \in \mathcal{A}$

   b. Define $\{A_i\}_{i\in\{1,2\}}$ by $A_1 = A \wedge A_2 = B$ then $A \bigcap B \underset{(1)}{=} \bigcap_{i\in K} A_i \in \mathcal{A}$

   c. As $B \in \mathcal{A}$ we have $X \setminus B \in \mathcal{A}$ so that $A \setminus B = A \bigcap B^c = A \bigcap X \bigcap B^c = A \bigcap (X\setminus B) \in \mathcal{A}$ (by 2.b)

3. As $X = X \setminus \emptyset$ we have by 18.6 (1,2) that $X \setminus \emptyset \in \mathcal{A}$ □

The following theorem shows how we can generate from a measurable space a new measurable space.

Theorem 18.9. *Let $\langle X, \mathcal{A}\rangle$ be a measurable space, $A \subseteq X$ then $\mathcal{A}_A = \{A \bigcap B | B \in \mathcal{A}\}$ is a σ-algebra on A making $\langle A, \mathcal{A}_A\rangle$ a measurable space. $\mathcal{A}_A$ is called the **sub-algebra** or **trace algebra** induced by A. Further $\mathcal{A}_X = \mathcal{A}$*

Proof. First as $\emptyset = A \bigcap \emptyset$ and $\emptyset \in \mathcal{A}$ we have that

$$\emptyset \in \mathcal{A}_A \tag{18.2}$$

Second if $B \in \mathcal{A}_A$ then $\exists E \in \mathcal{A}$ such that $B = A \bigcap E$ then

$$(X\setminus E) \bigcap A \underset{1.31}{=} (X \bigcap A) \setminus E$$
$$\underset{A\subseteq X}{=} A \setminus E$$

which as $(X \setminus E) \in \mathcal{A}$ proves that

$$A \setminus E \in \mathcal{A}_A \tag{18.3}$$

Finally if $\{A_i\}_{i\in\mathbb{N}} \subseteq \mathcal{A}_A$ then $\forall i \in \mathbb{N}$ there exists a $E_i \in \mathcal{A}$ such that $A_i = A \bigcap E_i$ hence

$$\bigcup_{i\in\mathbb{N}} A_i = \bigcup_{i\in\mathbb{N}} (A \bigcap E_i)$$
$$\underset{1.107}{=} A \bigcap \left(\bigcup_{i\in\mathbb{N}} E_i\right)$$

which as $\bigcup_{i\in\mathbb{N}} A_i \in \mathcal{A}$ proves that

$$\bigcup_{i\in\mathbb{N}} A_i \in \mathcal{A}_A \tag{18.4}$$

Finally as $\forall B \in \mathcal{A}$ we have $A \subseteq X$ so that $X \bigcap B = B$ proving that $\mathcal{A}_X = \{X \bigcap B | B \in \mathcal{A}\} = \{B | B \in \mathcal{A}\} = \mathcal{A}$ □

The following shows how to construct a algebra from a given set of subsets of a set and will be used in many cases to generate a σ-algebra.

Theorem 18.10. *Let X be a set and $\mathcal{A}\subseteq\mathcal{P}(X)$ then there exists a unique σ-algebra $\sigma[\mathcal{A}]$ on X such that*

1. $\mathcal{A}\subseteq\sigma[\mathcal{A}]$
2. *For every σ-algebra $\mathcal{B}$ with $\mathcal{A}\subseteq\mathcal{B}$ we have that $\sigma[\mathcal{A}]\subseteq\mathcal{B}$ [meaning that $\sigma[\mathcal{A}]$ is the smallest σ-algebra containing $\mathcal{A}$]*

Notation 18.11. *Elements of $\sigma[\mathcal{A}]$ are called Borel sets and $\sigma[\mathcal{A}]$ is called a* ***Borel algebra*** *and $\mathcal{A}$ is called the* ***generator*** *of the Borel algebra.*

Proof. Define $\mathcal{Q}=\{\mathcal{B}\subseteq\mathcal{P}(X)|\mathcal{B}$ is a σ-algebra and $\mathcal{A}\subseteq\mathcal{B}\}$ then as $\mathcal{P}(X)$ a σ-algebra (see 18.7) and $\mathcal{A}\subseteq\mathcal{P}(X)$ we have that $\mathcal{P}(X)\in\mathcal{Q}$ such that $\mathcal{Q}\neq\emptyset$. Define now $\sigma[\mathcal{A}]=\bigcap_{\mathcal{B}\in\mathcal{Q}}\mathcal{B}$ then as $\forall\mathcal{B}\in\mathcal{Q}$ we have $\mathcal{A}\subseteq\mathcal{B}\subseteq\mathcal{P}(X)$ it follows that

$$\mathcal{A}\subseteq\sigma[\mathcal{A}]\subseteq\mathcal{P}(X) \tag{18.5}$$

We prove now that $\sigma[\mathcal{A}]$ is a σ-algebra

1. $\forall\mathcal{B}$ we have that $\emptyset\in\mathcal{B}$ so that $\emptyset\in\bigcap_{\mathcal{B}\in\mathcal{Q}}\mathcal{B}$
2. If $A\in\sigma[\mathcal{A}]=\bigcap_{\mathcal{B}\in\mathcal{Q}}\mathcal{B}$ then $\forall\mathcal{B}\in\mathcal{Q}$ $A\in\mathcal{B}\underset{\mathcal{B}\text{ is a }\sigma\text{-algebra}}{\Rightarrow}X\setminus A\in\mathcal{B}$ proving that $X\setminus A\in\bigcap_{\mathcal{B}\in\mathcal{Q}}\mathcal{B}=\sigma[\mathcal{A}]$
3. If $\{A_i\}_{i\in\mathbb{N}}\subseteq\sigma[\mathcal{A}]$ then $\forall i\in\mathbb{N}$ we have $A_i\in\sigma[\mathcal{A}]$ so that $\forall\mathcal{B}\in\mathcal{Q}$ we have $A_i\in\mathcal{B}$, hence $\bigcup_{i\in\mathbb{N}}A_i\in\mathcal{B}$. So $\bigcup_{i\in\mathbb{N}}A_i\in\bigcap_{\mathcal{B}\in\mathcal{Q}}\mathcal{B}=\sigma[\mathcal{A}]$

hence

$$\sigma[\mathcal{A}]\text{ is a }\sigma\text{-algebra} \tag{18.6}$$

Further if there is a another σ-algebra $\mathcal{R}$ with $\mathcal{A}\subseteq\mathcal{R}$ then by definition $\mathcal{R}\in\mathcal{Q}$ so that $\sigma[\mathcal{A}]=\bigcap_{\mathcal{B}\in\mathcal{Q}}\mathcal{B}\subseteq\mathcal{R}$. This together with 18.5 and 18.6 proves the theorem. □

A special Borel algebra is the algebra generated by open sets in a topology.

Definition 18.12. *Let $\langle X,\mathcal{T}\rangle$ be a topological space then the* ***Borel σ-algebra*** *$\mathcal{B}[X,\mathcal{T}]$ is defined by $\mathcal{B}[X,\mathcal{T}]=\sigma[\mathcal{T}]$*

Notation 18.13. *We use the following notations for the following topological spaces*

1. *$\mathcal{B}[\mathbb{R}]$ is $\mathcal{B}[\mathbb{R},\mathcal{T}_{\mathbb{R}}]$ where $\mathcal{T}_{\mathbb{R}}$ is the topology on $\mathbb{R}$ generated by the norm $||$*
2. *$\mathcal{B}[\bar{\mathbb{R}}]$ is $\mathcal{B}[\bar{\mathbb{R}},\mathcal{T}_{\bar{\mathbb{R}}}]$ where $\mathcal{T}_{\bar{\mathbb{R}}}$ is the topology defined by the norm 17.36*
3. *Let $n\in\mathbb{N}$ then $\mathcal{B}[\mathbb{R}^n]$ is $\mathcal{B}[\mathbb{R}^n,\mathcal{T}_{\mathbb{R}^n}]$ is the product topology on $\mathbb{R}^n$ (see 12.80)*
4. *Let $n\in\mathbb{N}$ then $\mathcal{B}[\overline{\mathbb{R}^n}]$ is $\mathcal{B}[\bar{\mathbb{R}}^n,\mathcal{T}_{\bar{\mathbb{R}}^n}]$ is the box topology on $\bar{\mathbb{R}}^n$ (see 12.34)*

Actually we can generate the Borel σ-algebra generated by the topology on a set also by the set of closed sets as is proved in the following proposition.

Proposition 18.14. *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $\mathcal{C}=\{C\subseteq X|C\text{ is closed in }\mathcal{T}\}$ then $\mathcal{B}[X,\mathcal{T}]=\sigma[\mathcal{C}[$*

Proof. Let $C \in \mathcal{C}$ then there exist by definition a $U \in \mathcal{T}$ such that $C = X \setminus U \in \sigma[\mathcal{T}]$ [using the definition of a σ-algebra] proving that $\mathcal{C} \subseteq \boldsymbol{\sigma}[\mathcal{T}]$. As $\sigma[\mathcal{C}]$ is the smallest σ-algebra containing $\mathcal{C}$ we must have

$$\sigma[\mathcal{C}] \subseteq \sigma[\mathcal{T}] = \mathcal{B}[X, \mathcal{T}] \tag{18.7}$$

Further if $U \in \mathcal{T}$ then $X \setminus U$ is closed so that $X \setminus U \in \sigma[\mathcal{C}]$ [using the definition of a σ-algebra] proving that $\mathcal{T} \subseteq \sigma[\mathcal{C}]$. As $\sigma[\mathcal{T}]$ is the smallest σ-algebra containing $\mathcal{T}$ we have $\sigma[\mathcal{T}] \subseteq \sigma[\mathcal{C}]$ which by 18.7 proves that

$$\sigma[\mathcal{C}] = \sigma[\mathcal{T}] = \mathcal{B}[\mathcal{X}, \mathcal{T}]$$

□

18.1.2 Measure spaces

Now we know which sets we are going to work with we can introduce the concept of a measure to indicate the size of a set. In defining this concept we want that it behaves in the same intuitive ways as lengths, area's and volume's.

Definition 18.15. *A measure space is a triple $\langle X, \mathcal{A}, \mu \rangle$ where X is a set, $\mathcal{A}$ is a σ-algebra on X and μ is the graph of a function $\mu: X \rightarrow [0, \infty]$ such that*

1. $\mu(\emptyset) = 0$
2. *For every $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$ such that $\forall i, j \in \mathbb{N}$ with $i \neq j$ we have $A_i \bigcap A_j = \emptyset$ [a pairwise disjoint sequence of sets in $\mathcal{A}$] we have $\mu(\bigcup_{i \in \mathbb{N}} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$.*

the set function μ is called a ***measure****.*

Note that by 17.113 $\sum_{i=1}^{\infty} \mu(A_i) = \lim_{n \rightarrow \infty} \sum_{i=1}^{n} \mu(A_i) = \sup(\{\sum_{i=1}^{n} \mu(A_i) | n \in \mathbb{N}\})$. Also we have a more generalized definition of a infinite and finite sums of extended reals numbers (see 17.127) $\sum_{i \in I} x_i$, this allows us to have a alternative definition of measure spaces. First we need a little lemma.

Lemma 18.16. *Let I, J be sets and $\beta: I \rightarrow J$ a surjection, $\{A_i\}_{i \in J}$ a family of sets then $\bigcup_{i \in I} A_{\beta(i)} = \bigcup_{i \in J} A_i$*

Proof. First we prove that $\bigcup_{i \in I} A_{\beta(i)} \subseteq \bigcup_{i \in J} A_i$

$$\begin{aligned} x \in \bigcup_{i \in I} A_{\beta(i)} \quad &\Rightarrow \quad \exists i \in I \text{ such that } x \in A_{\beta(i)} \\ &\underset{\beta(i) \in J}{\Rightarrow} \quad x \in \bigcup_{i \in J} A_i \end{aligned}$$

Next we prove that $\bigcup_{i \in J} A_i \subseteq \bigcup_{i \in I} A_{\beta(i)}$

$$\begin{aligned} x \in \bigcup_{i \in J} A_i \quad &\Rightarrow \quad \exists i \in J \text{ such that } x \in A_i \\ &\underset{\beta \text{ is surjective}}{\Rightarrow} \quad \exists j \in I \text{ such that } i = \beta(j) \\ &\Rightarrow \quad x \in A_{\beta(j)} \\ &\Rightarrow \quad \bigcup_{i \in I} A_{\beta(i)} \end{aligned}$$

□

Proposition 18.17. *Let X be a set, $\mathcal{A}$ a σ-algebra then $\langle X, \mathcal{A}, \mu\rangle$ is a measure space iff*

1. $\mu(\emptyset) = 0$
2. *For every $\{A_i\}_{k \in K} \subseteq \mathcal{A}$ such that K is countable and $\forall i, j \in K$ with $i \neq j$ we have $A_i \bigcap A_j = \emptyset$ we have $\mu(\bigcup_{i \in K} A_i) = \sum_{i \in K} \mu(A_i)$*

Proof. It is enough to prove that (2) of the proposition is equivalent to (2) of the definition

$\Rightarrow$. Let $\{A_i\}_{i \in K} \subseteq \mathcal{A}$ such that $\forall i, j \in K$ with $i \neq j$ we have $A_i \bigcap A_j = \emptyset$ then we have the following cases to consider for K:

$K = \emptyset$. then $\mu(\bigcup_{i \in K} A_i) \underset{1.100}{=} \mu(\emptyset) = 0 \underset{17.127}{=} \sum_{i \in K} \mu(A_i)$

K is finite non empty. then there exists a $n \in \mathbb{N}$ such that $\beta: \{1, \ldots, n\} \to K$ define then $\{B_i\}_{i \in \mathbb{N}}$ by $B_i = \begin{cases} A_{\beta(i)} \text{ if } i \in \{1, \ldots, n\} \\ \emptyset \text{ if } i \in \mathbb{N} \setminus \{1, \ldots, n\} \end{cases}$ then $\{B_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$ and if $i, j \in \mathbb{N}$ with $i \neq j$ we have either

$i, j \in \{1, \ldots, n\}$. then $B_i \bigcap B_j = A_{\beta(i)} \bigcap A_{\beta(j)} \underset{\beta \text{ is a bijection} \Rightarrow \beta(i) \neq \beta(j)}{=} \emptyset$

$1 < i \wedge j \in \{1, \ldots, n\}$. then $B_i \bigcap B_j = \emptyset \bigcap B_j = \emptyset$

$i \in \{1, \ldots, n\} \wedge 1 < j$. then $B_i \bigcap B_j = B_i \bigcap B_j = \emptyset$

$1 < i \wedge 1 < j$. then $B_i \bigcap B_j = \emptyset \bigcap \emptyset = \emptyset$

proving that $\{B_i\}_{i \in \mathbb{N}}$ is pairwise disjoint so that

$$\mu\left(\bigcup_{i \in \mathbb{N}} B_i\right) = \sum_{i=1}^{\infty} \mu(B_i) \tag{18.8}$$

Now $\forall i \in \mathbb{N} \setminus \{1, \ldots, n\}$ we have $\mu(B_i) = \mu(\emptyset) = 0$ so that

$$\begin{aligned} \sum_{i=1}^{\infty} \mu(B_i) &\underset{17.128}{=} \sum_{i \in \mathbb{N}} \mu(B_i) \\ &\underset{17.130}{=} \sum_{i \in \{1, \ldots, n\}} \mu(B_i) \\ &= \sum_{i \in \{1, \ldots, n\}} \mu(A_{\beta(i)}) \\ &\underset{10.44}{=} \sum_{i=1}^{n} \mu(A_{\beta(i)}) \\ &= \sum_{i \in K} \mu(A_i) \end{aligned} \tag{18.9}$$

Further

$$\begin{aligned}\bigcup_{i\in K} A_i &= \bigcup_{i\in\{1,\ldots,n\}} A_{\beta(i)}\\ &= \bigcup_{i\in\{1,\ldots n\}} B_i\\ &= \bigcup_{i\in\mathbb{N}} B_i\end{aligned}$$

so that using 18.8 and 18.9 we have

$$\mu\left(\bigcup_{i\in K} A_i\right)=\sum_{i\in K} \mu(A_i)$$

***K* is infinite countable.** then there exists a $\beta\colon\mathbb{N}\rightarrow K$ then and if i, $j\in\mathbb{N}$ with $i\neq j$ we have as β is a bijection that $\beta(i)\neq\beta(j)$ so that $A_{\beta(i)}\bigcap A_{\beta(j)}=\emptyset$ proving that $\{A_{\beta(i)}\}_{i\in\mathbb{N}}$ is pairwise disjoint

$$\begin{aligned}\mu\left(\bigcup_{i\in K} A_i\right) &\underset{18.16}{=} \mu\left(\bigcup_{i\in\mathbb{N}} A_{\beta(i)}\right)\\ &= \sum_{i=1}^{\infty} \mu(A_{\beta(i)})\\ &\underset{17.127}{=} \sum_{i\in K} \mu(A_i)\end{aligned}$$

So in all cases we have that

$$\mu\left(\bigcup_{i\in K} A_i\right)=\sum_{i\in K} \mu(A_i)$$

⇐. Let $\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{A}$ be a family of pairwise disjoint then as $\mathbb{N}$ is countable we have

$$\begin{aligned}\mu\left(\bigcup_{i\in\mathbb{N}} A_i\right) &= \sum_{i\in\mathbb{N}} \mu(A_i)\\ &\underset{17.128}{=} \sum_{i=1}^{\infty} \mu(A_i)\end{aligned}$$

□

In this treatise we will encounter many times that the sentence let $\{A_i\}_{i\in I}$ be a family of pairwise disjoint sets and then work with the union $\bigcup_{i\in I} A_i$. To avoid this excessive notation we introduce the following notations.

Notation 18.18. *If we define $\{A_i\}$ and then refer to $\bigsqcup_{i\in I} A_i$ we mean that $\{A_i\}_{i\in I}$ is pairwise disjoint and that $\bigsqcup_{i\in I} A_i=\bigcup_{i\in I} A_i$.*

Given the role that pairwise disjoint sequences of sets play in measure theory, it is useful to turn a sequence of sets in a pairwise disjoint union of sets. The following lemma can be used to do this.

Lemma 18.19. *Let $\{A_i\}_{i\in\mathbb{N}}$ is a sequence of sets then if we define $\{B_i\}_{i\in\mathbb{N}}$ by $B_i=\begin{cases} A_1 \text{ if } i=1 \\ A_i\setminus(\bigcup_{j\in\{1,\ldots,i-1\}} A_j) \text{ if } i\in\mathbb{N}\setminus\{1\}\end{cases}$ we have that $\bigcup_{i\in\mathbb{N}} A_i=\bigsqcup_{i\in\mathbb{N}} B_i$. Furthermore if $\langle X,\mathcal{A}\rangle$ is a measurable space and $\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{A}$ then $\{B_i\}_{i\in\mathbb{N}}\subseteq\mathcal{A}$*

Proof. First we prove that $\{B_i\}_{i\in\mathbb{N}}$ is pairwise disjoint. So let $i,j\in\mathbb{N}$ with $i\neq j$ then we may assume that $i<j$ [otherwise interchange i and j] so that $1\leqslant i<j$ and thus that $j\neq 1$. Assume that $x\in B_i\bigcap B_j$ then $x\in B_i$ and $x\in B_j$. If $i=1$ then $B_i=A_1$ so that $x\in A_1=A_i$ and if $1<i$ then $B_i=A_i\setminus\left(\bigcup_{k\in\{1,\ldots,i-1\}} A_k\right)$ so that $x\in A_i$, hence we must have $x\in A_i$. As $i<j$ we have $i\leqslant j-1$ hence $x\in\bigcup_{k\in\{1,\ldots,j-1\}} A_k$ so that $x\notin A_j\setminus\left(\bigcup_{k\in\{1,\ldots,j-1\}} A_k\right)$ a contradiction. So we must have that $B_i\bigcap B_j$. This proves that $\{B_i\}_{i\in\mathbb{N}}$ is pairwise disjoint.

Second we have as $\forall i\in\mathbb{N}$ $B_i\subseteq A_i$ we have that

$$\bigcup_{i\in\mathbb{N}} B_i\subseteq\bigcup_{i\in\mathbb{N}} A_i. \tag{18.10}$$

For the opposite inclusion let $x\in\bigcup_{i\in\mathbb{N}} A_i$ then there exists a $i\in\mathbb{N}$ such that $x\in A_i$, so that $I=\{j\in\{1,\ldots,i\}|x\in A_j\}\neq\emptyset$ hence $m=\min(I)$ exists. We have two cases for m to consider

$m=1$. then $x\in A_1=B_1\subseteq\bigcup_{i\in\mathbb{N}} B_i$

$1<m$. then by definition of the minimum we must have that $x\in A_m$ and $\forall i\in\{1,\ldots,m-1\}$ [so that $i<m$] we must have that $x\notin A_i$ hence $x\notin\bigcup_{i\in\{1,\ldots,m-1\}} A_i$. Hence $x\in A_m\setminus\left(\bigcup_{i\in\{1,\ldots,m-1\}} A_i\right)=B_m\subseteq\bigcup_{i\in\mathbb{N}} B_i$

So we must have that $\bigcup_{i\in\mathbb{N}} A_i\subseteq\bigcup_{i\in\mathbb{N}} B_i$ which using 18.10 and the pairwise disjointedness gives

$$\bigcup_{i\in\mathbb{N}} A_i=\bigsqcup_{i\in\mathbb{N}} B_i$$

Finally if $\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{A}$ then given $i\in\mathbb{N}$ we have either

$i=1$. then $B_i=B_1=A_1\in\mathcal{A}$

$1<i$. then $B_i=A_i\setminus\left(\bigcup_{j\in\{1,\ldots,i-1\}} A_j\right)$. By 18.6 (2) we have that $\bigcup_{j\in\{1,\ldots,i-1\}} A_j\in\mathcal{A}$ and as $A_i\in\mathcal{A}$ we can use 18.8 proving that $B_i=A_i\setminus\left(\bigcup_{j\in\{1,\ldots,i-1\}} A_j\right)\in\mathcal{A}$ □

We have the following properties of a measure space

Proposition 18.20. *Let $\langle X,\mathcal{A},\mu\rangle$ be a measure space then we have*

1. *$\forall A,B\in\mathcal{A}$ with $A\bigcap B=\emptyset$ we have $\mu(A\bigcup B)=\mu(A)+\mu(B)$*

2. *$\forall A, B \in \mathcal{A}$ with $A \subseteq B$ we have $\mu(A) \leqslant \mu(B)$*

3. *$\forall A, B \in \mathcal{A}$ we have $\mu(A \bigcup B) \leqslant \mu(A) + \mu(B)$*

4. *$\forall \{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$ we have $\mu(\bigcup_{i \in \mathbb{N}} A_i) \leqslant \sum_{i \in \mathbb{N}} \mu(A_i)$*

5. *If $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$ is such that $\forall i \in \mathbb{N}$ we have $A_i \subseteq A_{i+1}$ [$\{A_i\}_{i \in \mathbb{N}}$ is a increasing sequence] then*

$$\mu\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \lim_{i \to \infty} \mu(E_i) = \sup(\{\mu(E_i) | i \in \mathbb{N}\})$$

6. *If $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$ is such that $\forall i \in \mathbb{N}$ we have that $A_{i+1} \subseteq A_i$ [$\{A_i\}_{i \in \mathbb{N}}$ is a decreasing sequence] and $\exists k \in \mathbb{N}$ such that $\mu(A_k) < \infty$ then*

$$\mu\left(\bigcap_{i \in \mathbb{N}} A_i\right) = \lim_{n \to \infty} \mu(A_n) = \inf(\{\mu(A_n) | n \in \mathbb{N}\})$$

7. *Let $A, B \in \mathcal{A}$ such that $\mu(B) < \mu(A)$ then $0 < \mu(A \setminus B)$*

Proof.

1. Define $\{A_i\}_{i \in \{1,2\}}$ by $A_1 = A$ and $A_2 = B$ then as $\{1,2\}$ is countable we have by 18.17 (2) that $\mu(A \bigcup B) = \mu(\bigcup_{i \in \{1,2\}} A_i) = \sum_{i \in \{1,.2\}} \mu(A_i) = \sum_{i=1}^{2} \mu(A_i) = \mu(A_1) + \mu(A_2) = \mu(A) + \mu(B)$

2. If $A \subseteq B$ then $B \underset{1.31}{=} (B \setminus A) \bigcup A$ with $(B \setminus A) \bigcap A = \emptyset$ hence by (1) we have that $\mu(A) = \mu(B \setminus A) + \mu(A) \geqslant \mu(A)$ as $\mu(B \setminus A) \geqslant 0$

3. As $(A \setminus B) \bigcup B \underset{1.31}{=} A \bigcup B$ and $(A \setminus B) \bigcap B$ so by (1) we have $\mu(A \bigcup B) = \mu(A \setminus B) + \mu(B) \leqslant \mu(A) + \mu(B)$ [using $A \setminus B \subseteq A$ and (2)]

4. Define $\{B_i\}_{i \in \mathbb{N}}$ by $B_i = \begin{cases} A_1 \text{ if } i = 1 \\ A_i \setminus (\bigcup_{j \in \{1, \ldots, i-1\}} A_j) \end{cases}$ then using 18.19 we have that $\bigcup_{i \in \mathbb{N}} A_i = \bigsqcup_{i \in \mathbb{N}} B_i$ and $\{B_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$, further we have $B_i \subseteq A_i \underset{(2)}{\Rightarrow} \mu(B_i) \leqslant \mu(A_i)$. So $\mu(\bigcup_{i \in \mathbb{N}} A_i) = \mu(\bigsqcup_{i \in \mathbb{N}} B_i) \underset{\mu \text{ is a measure}}{=} \sum_{i=1}^{\infty} \mu(B_i) \leqslant_{17.115} \sum_{i=1}^{\infty} \mu(A_i)$.

5. Define $\{B_i\}_{i \in \mathbb{N}}$ by $B_i = \begin{cases} A_1 \text{ if } i = 1 \\ A_i \setminus A_{i-1} \text{ if } i \in \mathbb{N} \setminus \{1\} \end{cases}$ then using 18.8 we have that $\{B_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$. We prove now that

$$\bigcup_{i \in \mathbb{N}} A_i = \bigsqcup_{i \in \mathbb{N}} B_i \tag{18.11}$$

Proof. Let $i, j \in \mathbb{N}$ be such that $i \neq j$ then we may assume that $i < j$ [otherwise interchange i and j]. Assume that $x \in B_i \bigcap B_j$ then $x \in B_i$ and $x \in B_j$. Using 5.77 we have that $A_i \subseteq A_{j-1}$ so as $x \in B_i \subseteq A_i$ we have $x \in A_{j-1}$ and thus $x \notin A_j \setminus A_{j-1} = B_j$ a contradiction. Hence we have that $\{B_i\}_{i \in \mathbb{N}}$ is pairwise disjoint.

As $\forall i \in \mathbb{N}$ we have $B_i \subseteq A_i$ so that $\bigcup_{i \in \mathbb{N}} B_i \subseteq \bigcup_{i \in \mathbb{N}} A_i$. For the opposite inequality take $x \in \bigcup_{i \in \mathbb{N}} A_i$ then there exists a $i \in \mathbb{N}$ such that $x \in A_i$ so that $I = \{j \in \{1, \ldots, i\} | x \in A_j\} \neq \emptyset$ hence $m = \min(I)$ exists. We have two cases to consider for m

$\boldsymbol{m = 1.}$ then $x \in A_1 = B_1 \subseteq \bigcup_{i \in \mathbb{N}} B_i$

$\boldsymbol{1 < m.}$ then $x \in A_m$ and $x \notin A_{m-1}$ so that $x \in B_m \subseteq \bigcup_{i \in \mathbb{N}} B_i$

proving that $\bigcup_{i \in \mathbb{N}} A_i = \bigsqcup_{i \in \mathbb{N}} B_i$ □

Next we prove by induction that $\forall n \in \mathbb{N}$

$$\mu(A_n) = \sum_{j=1}^{n} \mu(B_j) \tag{18.12}$$

Proof. So let $\mathcal{S} = \{n \in \mathbb{N} | \mu(A_n) = \sum_{j=1}^{n} \mu(B_j)\}$ then

$\boldsymbol{1 \in \mathcal{S}.}$ for $n = 1$ we have $\mu(A_1) = \mu(B_1) = \sum_{i=1}^{1} \mu(B_i)$ so that $1 \in \mathcal{S}$

$\boldsymbol{n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}.}$ let $n \in \mathcal{S}$ then $B_{n+1} = A_{n+1} \setminus A_n$ so that as $A_n \subseteq A_{n+1}$ we have that $A_{n+1} = A_n \bigsqcup (A_{n+1} \setminus A_n) = A_n \bigsqcup B_{n+1}$. Hence

$$\begin{aligned}
\mu(A_{n+1}) &= \mu(A_n) + \mu(B_{n+1}) \\
&\underset{n \in \mathcal{S}}{=} \sum_{i=1}^{n} \mu(B_i) + \mu(B_{n+1}) \\
&= \sum_{i=1}^{n+1} \mu(B_i)
\end{aligned}$$

proving that $n + 1 \in \mathcal{S}$ □

Next as $\forall i \in \mathbb{N}$ we have $A_i \subseteq A_1$ it follows that

$$\forall i \in \mathbb{N} \text{ we have } \mu(A_i) \leqslant \mu(A_{i+1}) \tag{18.13}$$

$$\begin{aligned}
\mu\left(\bigcup_{i \in \mathbb{N}} A_i\right) &= \mu\left(\bigsqcup_{i \in \mathbb{N}} B_i\right) \\
&= \sum_{i=1}^{\infty} \mu(B_i) \\
&= \lim_{n \to \infty} \sum_{i=1}^{n} \mu(B_i) \\
&\underset{18.12}{=} \lim_{n \to \infty} \mu(A_n) \\
&\underset{18.13 \text{ and } 17.83}{=} \sup(\{\mu(A_n) | n \in \mathbb{N}\})
\end{aligned}$$

proving (5).

6. Let $k\in\mathbb{N}$ be such that $\mu(A_k)<\infty$ and define $\{B_i\}_{i\in\mathbb{N}}$ by $A_k\setminus A_i$. Then

$$\forall i\in\mathbb{N}\text{ we have }B_i\in\mathcal{A} \text{ and } B_i\subseteq B_{i+1} \tag{18.14}$$

Proof. First $B_i=A_k\setminus A_{k+i}\in\mathcal{A}$ because A_k, A_{k+i} and 18.8. If $x\in B_i=A_k\setminus A_{k+i}$ then $x\in A_k$ and $x\notin A_{k+i}\subseteq_{\text{hypothesis and 5.77}} A_k$ so that $x\in A_k$ and $x\notin A_{i+1}$ proving that $x\in B_{i+1}$ □

Using (5) proves

$$\mu\left(\bigcup_{i\in\mathbb{N}} B_i\right)=\lim_{i\rightarrow\infty}\mu(B_i) \tag{18.15}$$

Further as $A_{k+i}\subseteq A_k$ we have that $A_k\underset{1.31}{=}(A_k\setminus A_{k+i})\bigsqcup A_{k+i}=B_i\bigsqcup A_{k+i}$ so that $\mu(A_k)=\mu(B_i)+\mu(A_{k+i})$ and as $\mu(A_k)<\infty$ we have $\mu(A_{k+i})\leqslant\mu(A_k)<\infty$ so that $\mu(B_i)=\mu(A_k)-\mu(A_{k+i})$. Substituting this in 18.15 gives

$$\begin{aligned}\mu\left(\bigcup_{i\in\mathbb{N}} B_i\right) &= \lim_{i\rightarrow\infty}(\mu(A_k)-\mu(A_{k+i}))\\ &\underset{17.81 \text{ and } 17.89}{=} \mu(A_k)-\lim_{i\rightarrow\infty}\mu(A_{k+i})\\ &\underset{17.80}{=} \mu(A_k)-\lim_{i\rightarrow\infty}\mu(A_i)\end{aligned} \tag{18.16}$$

Now as $\forall i\in\mathbb{N}$ we have $B_i=A_k\setminus A_{k+i}\subseteq A_k$ it follows that $\bigcup_{i\in\mathbb{N}}B_i\subseteq A_k$ hence we have

$$A_k=\left(A_k\setminus\left(\bigcup_{i\in\mathbb{N}} B_i\right)\right)\bigsqcup\left(\bigcup_{i\in\mathbb{N}} B_i\right) \tag{18.17}$$

As

$$\begin{aligned}A_k\setminus\left(\bigcup_{i\in\mathbb{N}} B_i\right) &\underset{1.108}{=} \bigcap_{i\in\mathbb{N}}(A_k\setminus B_i)\\ &= \bigcap_{i\in\mathbb{N}}(A_k\setminus(A_k\setminus A_{k+i}))\\ &\underset{1.31}{=} \bigcap_{i\in\mathbb{N}} A_{k+i}\end{aligned} \tag{18.18}$$

Now we have

$$\begin{aligned}x\in\bigcap_{i\in\mathbb{N}}A_i &\Rightarrow \forall i\in\mathbb{N}\text{ we have }x\in A_i\\ &\underset{k+i\in\mathbb{N}}{\Rightarrow} \forall i\in\mathbb{N}\text{ we have }x\in A_{i+k}\\ &\Rightarrow x\in\bigcap_{i\in\mathbb{N}}A_{i+k}\end{aligned}$$

proving that

$$\bigcap_{i\in\mathbb{N}}A_i\subseteq\bigcap_{i\in\mathbb{N}}A_{i+k} \tag{18.19}$$

For the opposite inclusion let $x \in \bigcap_{i \in \mathbb{N}} A_{i+k}$ then $\forall i \in \mathbb{N}$ we have $x \in A_{i+k}$. Take $j \in \mathbb{N}$ then $x \in A_{j+k} \subseteq A_j$ proving that $x \in \bigcap_{j \in \mathbb{N}} A_j$ proving that $\bigcap_{i \in \mathbb{N}} A_{i+k} \subseteq \bigcap_{i \in \mathbb{N}} A_i$. Using 18.19 we have then that $\bigcap_{i \in \mathbb{N}} A_{i+k} = \bigcap_{i \in \mathbb{N}} A_i$, substituting this in 18.18 gives $A_k \setminus (\bigcup_{i \in \mathbb{N}} B_i) = \bigcap_{i \in \mathbb{N}} A_i$ and then using 18.17 we have

$$A_k = \left(\bigcap_{i \in \mathbb{N}} A_i\right) \bigsqcup \left(\bigcup_{i \in \mathbb{N}} B_i\right)$$

so that

$$\begin{aligned} \mu(A_k) &= \mu\left(\bigcap_{i \in \mathbb{N}} A_i\right) + \mu\left(\bigcup_{i \in \mathbb{N}} B_i\right) \\ &\underset{18.16}{=} \mu\left(\bigcap_{i \in \mathbb{N}} A_i\right) + \mu(A_k) - \lim_{i \to \infty} \mu(A_i) \end{aligned}$$

or as $\mu(A_k) < \infty$ so that also $\mu(\bigcap_{i \in \mathbb{N}} A_i) < 0$ [because $\bigcap_{i \in \mathbb{N}} A_i \subseteq A_k$ we have $-\mu(\bigcap_{i \in \mathbb{N}} A_i) = -\lim_{i \to \infty} \mu(A_i)$ proving that

$$\mu\left(\bigcap_{i \in \mathbb{N}} A_i\right) = \lim_{i \to \infty} \mu(A_i) \underset{17.83}{=} \inf(\{\mu(A_i) | i \in \mathbb{N}\})$$

7. As $\mu(B) < \mu(A)$ we have that $\mu(A \bigcap B) < \mu(A) \leqslant \infty$ so that also $\mu(A \bigcap B) < \infty$. Further as $A = (A \setminus B) \bigsqcup (A \bigcap B)$ we have $\mu(A \setminus B) + \mu(A \bigcap B) = \mu(A) > \mu(A \bigcap B)$ and as $\mu(A \bigcap B) < \infty$ it follows that $0 < \mu(A \setminus B)$ □

If we have a measure space then we can induce a algebra and a measure on a other space using a function between the two spaces.

Proposition 18.21. *Let $\langle X, \mathcal{A}, \mu \rangle$ be a measure space, Y a set and $f: X \to Y$ then $\langle Y, \mathcal{A}_f, \mu_{\langle f^{-1} \rangle} \rangle$ is a measure space where $\mathcal{A}_f = \{B \in \mathcal{P}(Y) | f^{-1}(B) \in \mathcal{A}\}$ and $\mu_{\langle f^{-1} \rangle}: \mathcal{A}_f \to [0, \infty]$ is defined by $\mu_{\langle f^{-1} \rangle}(B) = \mu(f^{-1}(B))$. $\mu_{\langle f^{-1} \rangle}$ is called the **image measure** on Y.*

Proof. First we prove that $\mathcal{A}_f$ is a σ-algebra on Y

1. As $f^{-1}(\emptyset) = \emptyset \in \mathcal{A}$ so that $\emptyset \in \mathcal{A}_f$
2. Let $B \in \mathcal{A}_f$ then $f^{-1}(B) \in \mathcal{A}$ so that $X \setminus f^{-1}(B) \in \mathcal{A}$. Now $f^{-1}(Y \setminus B) \underset{2.54\ (4)}{=} X \setminus f^{-1}(B) \in \mathcal{A}$ proving that $X \setminus B \in \mathcal{A}_f$
3. Let $\{B_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}_f$ then $\forall i \in \mathbb{N}$ we have $f^{-1}(B) \in \mathcal{A}$. Now $f^{-1}(\bigcup_{i \in \mathbb{N}} B_i) \underset{2.58}{=} \bigcup_{i \in \mathbb{N}} f^{-1}(B_i) \in \mathcal{A}$ proving that $\bigcup_{i \in \mathbb{N}} B_i \in \mathcal{A}$.

Next we must prove that $\mu_{\langle f^{-1} \rangle}$ is a measure on $\mathcal{A}_f$.

1. $\mu_{\langle f^{-1} \rangle}(\emptyset) = \mu(f^{-1}(\emptyset)) = \mu(\emptyset) = 0$

2. Let $\{B_i\}_{i\in\mathbb{N}}\subseteq\mathcal{A}_f$ such that $\forall i,j\in\mathcal{A}$ we have $B_i\bigcap B_j=\emptyset$. As $f^{-1}(B_i)\bigcap f^{-1}(B_j)\underset{2.58}{=}f^{-1}(B_i\bigcap B_j)=f^{-1}(\emptyset)=0$ proving that $\{f^{-1}(A_i)\}_{i\in\mathbb{N}}\subseteq\mathcal{A}$ is pairwise disjoint. So
$$\begin{aligned}\mu_{\langle f^{-1}\rangle}\left(\bigsqcup_{i\in\mathbb{N}}B_i\right) &\underset{\text{def}}{=} \mu\left(f^{-1}\left(\bigcup_{i\in\mathbb{N}}B_i\right)\right)\\ &\underset{2.58}{=} \mu\left(\bigsqcup_{i\in\mathbb{N}}f^{-1}(B_i)\right)\\ &= \sum_{i\in\mathbb{N}}\mu(f^{-1}(B_i))\\ &= \sum_{i\in\mathbb{N}}\mu_{\langle f^{-1}\rangle}(B_i)\end{aligned}$$
□

18.1.3 Negligible sets

Definition 18.22. (Negligible Set) *Let $\langle X,\mathcal{A},\mu\rangle$ be a measure space and $A\subseteq X$, then A is μ-**negligible** iff $\exists B\in\mathcal{A}$ such that $\mu(B)=0$ and $A\subseteq B$. The set of all the μ-negligible sets is noted by $\mathcal{N}_\mu$ so $\mathcal{N}_\mu=\{A\in\mathcal{P}(X)|A\,\text{is}\,\mu\text{-negligible}\}$*

Note 18.23. Note that it is not required that a set is part of the σ-algebra to be negligible.

We have the following properties for $\mathcal{N}_\mu$

Proposition 18.24. *Let $\langle X,\mathcal{A},\mu\rangle$ be a measure space then we have*

1. $\emptyset\in\mathcal{N}_\mu$
2. *If $A\in\mathcal{N}_\mu$ then if $B\subseteq A$ we have $B\in\mathcal{N}_\mu$*
3. *$\forall\{A_i\}_{i\in K}\subseteq\mathcal{N}_\mu$ where K is countable then $\bigcup_{i\in K}N_i\in\mathcal{N}_\mu$*

Proof.

1. *This is trivial as $\emptyset\subseteq\emptyset$ and $\mu(\emptyset)=0$*
2. *As $A\in\mathcal{N}$ there exists a $A'\in\mathcal{A}$ such that $B\subseteq A\subseteq A'$ and $\mu(A')=0$ proving that $B\in\mathcal{N}_\mu$*
3. *$\forall i\in K$ there exists a $N_i\in\mathcal{A}$ such that $A_i\subseteq N_i$ and $\mu(N_i)$ then $\bigcup_{i\in K}A_i\subseteq\bigcup_{i\in K}N_i$ and $\mu(\bigcup_{i\in K}N_i)\leqslant_{18.20}\sum_{i\in K}\mu(N_i)=\sum_{i\in K}0=0$* □

Definition 18.25. (Conegligible Set) *Let $\langle X,\mathcal{A},\mu\rangle$ be a measure space then $A\in\mathcal{P}(X)$ is a μ-conegligible set if $X\setminus A$ is μ-negligible set.*

Note 18.26. If from the context it is clear which measure μ is used then we us the term conegligible instead of μ-conegligible.

Conegligible sets have the following properties.

Proposition 18.27. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space then we have*

1. *X is μ-conegligible.*
2. *If A is μ-conegligible and $B \in \mathcal{P}(A)$ such that $A \subseteq B$ then B is μ-conegligible.*
3. *If A, B are μ-conegligible then $A \bigcup B$ is μ-conegligible*
4. *Let $\{A_i\}_{i \in K}$ be a countable family of μ-conegligible sets $\bigcap_{i \in K} A_i$ is μ-conegligible.*
5. *If A is μ-conegligible then there exists a $B \in \mathcal{A}$ such that $B \subseteq A$ and B is μ-conegligible.*
6. *If A is μ-conegligible and B is μ-negligible then $A \setminus B$ is μ-conegligible*
7. *Let $A, B \subseteq X$ such that $A \bigcup B$ is μ-conegligible and B is μ-negligible then A is μ-conegligible.*

Proof.

1. As $X \setminus X = \emptyset$ we using 18.24 that $X \setminus X$ is μ-negligible so that X is μ-conegligible
2. As $A \subseteq B$ we have that $X \setminus B \subseteq X \setminus A$ and as A is μ-conegligible so that there exists a $N \in \mathcal{A}$ with $\mu(N) = 0$ and $X \setminus B \subseteq X \setminus A \subseteq N$ proving that B is μ-conegligible.
3. This follows from $A \subseteq A \bigcup B$ and (2)
4. As $\forall i \in K$ A_i is μ-conegligible there exists $N_i \in \mathcal{A}$ such that $X \setminus A_i \subseteq N_i$ and $\mu(N_i) = 0$. Hence $X \setminus (\bigcap_{i \in K} A_i) = \bigcup_{i \in K} (X \setminus A_i) \subseteq \bigcup_{i \in K} N_i$ and $\mu(\bigcup_{i \in K} N_i) \leqslant_{18.20} \sum_{i \in K} \mu(N_i) = \sum_{i \in K} 0 = 0$ proving that $\bigcap_{i \in K} A_i$ is μ-conegligible.
5. As A is μ-conegligible then there exists a $N \in \mathcal{A}$ with $\mu(N) = 0$ so that $X \setminus A \subseteq N$. Take $B = X \setminus N$ then using the definition of a σ-algebra we have that $B \in \mathcal{A}$, also as $B = X \setminus N \subseteq (X \setminus (X \setminus A)) = A$ we have that $B \subseteq A$, finally as $X \setminus B = X \setminus (X \setminus N) = N$ and $\mu(N) = 0$ we have that B is μ-conegligible.
6. As A is μ-conegligible we have that $X \setminus A$ is μ-negligible. Next $X \setminus (A \setminus B) \underset{1.31}{=} (X \setminus A) \bigcup B$ which μ-negligible (see 18.24) hence $A \setminus B$ is conegligible.
7. If $x \in (A \bigcup B) \setminus (B \setminus A)$ then $x \notin B \setminus A$ and either (not exclusive)

   ***x* ∈ *A*.** so that trivially $x \in A$

   ***x* ∈ *B*.** then if $x \notin A$ we have $x \in B \setminus A$ contradicting $x \notin B \setminus A$ so we must have that $x \in A$

   proving that

   $$(A \bigcup B) \setminus (B \setminus A) \subseteq A$$

Further if $x \in A$ then $x \in A \bigcup B$ and $x \notin B \setminus A$ so that $x \in (A \bigcup B) \setminus (B \setminus A)$ which by the above proves that

$$A = \left(A \bigcup B\right) \setminus (B \setminus A)$$

As B is μ-negligible we have as $B \setminus A \subseteq B$ that by 18.24 (2) $B \setminus A$ is μ-negligible so using (6) it follows that A is μ-conegligible.

□

We use conegligible set in the term **almost every predicates** as expressed in the following notation

Notation 18.28. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space and $P(x)$ a assertion about $x \in X$ then $P(x)$ for* ***μ almost every*** *$x \in X$ iff $\{x \in X | P(x)\}$ is μ-conegligible. Other notations are*

$$P(x)\, a.e.\, [\mu]$$

To avoid extra notation we omit the μ if it is clear from the context what the measure involved is. So then we say that **$P(x)$ for almost every** *$x \in X$ or write*

$$P(X)\, a.e.\, [\mu]$$

The following is a list of **almost every** predicates used in this document

Definition 18.29. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measurable set then for partial functions f, $g: X \to \mathbb{R}$ with $\mathrm{dom}(f), \mathrm{dom}(g) \subseteq X$ we have*

1. $f \leqslant_{a.e.[\mu]} g$ *iff* $\{x \in \mathrm{dom}(f) \bigcap \mathrm{dom}(g) | f(x) \leqslant g(x)\}$ *is μ-conegligible*
2. $f <_{a.e[\mu]} g$ *iff* $\{x \in \mathrm{dom}(f) \bigcap \mathrm{dom}(g) | f(x) < g(x)\}$ *is μ-conegligible*
3. $f =_{a.e.[\mu]} g$ *iff* $\{x \in \mathrm{dom}(f) \bigcap \mathrm{dom}(g) | f(x) = g(x)\}$ *is μ-conegligible*
4. *Take $a \in \mathbb{R}$ then $a \leqslant_{a.e.[\mu]} f$ iff* $\{x \in \mathrm{dom}(f) | a \leqslant f(x)\}$ *is μ-conegligible*
5. *Take $a \in \mathbb{R}$ then $a <_{a.e.[\mu]} f$ iff* $\{x \in \mathrm{dom}(f) | a < f(x)\}$ *is μ-conegligible*
6. *Take $a \in \mathbb{R}$ then $f \leqslant_{a.e.[\mu]} a$ iff* $\{x \in \mathrm{dom}(f) | f(x) \leqslant a\}$ *is μ-conegligible*
7. *Take $a \in \mathbb{R}$ then $f <_{a.e.[\mu]} a$ iff* $\{x \in \mathrm{dom}(f) | f(x) < a\}$ *is μ-conegligible*
8. *Take $a \in \mathbb{R}$ then $f =_{a.e.[\mu]} a$ iff* $\{x \in \mathrm{dom}(f) | f(x) = a\}$ *is μ-conegligible*
9. *Take $a \in \mathbb{R}$ then $f \neq_{a.e.[\mu]} a$ iff* $\{x \in \mathrm{dom}(f) | f(x) \neq a\}$ *is μ-conegligible*

If $\mathrm{dom}(f) = X$ [so f is a function] we have $\mathrm{dom}(f) \bigcap \mathrm{dom}(g) = \mathrm{dom}(g)$ then

1. $f \leqslant_{a.e.[\mu]} g$ *iff* $\{x \in \mathrm{dom}(g) | f(x) \leqslant g(x)\}$ *is μ-conegligible*
2. $f <_{a.e[\mu]} g$ *iff* $\{x \in \mathrm{dom}(g) | f(x) < g(x)\}$ *is μ-conegligible*
3. $g \leqslant_{a.e.[\mu]} f$ *iff* $\{x \in \mathrm{dom}(g) | g(x) \leqslant f(x)\}$ *is μ-conegligible*
4. $g <_{a.e.[\mu]} f$ *iff* $\{x \in \mathrm{dom}(g) | g(x) < f(x)\}$ *is μ-conegligible*
5. $f =_{a.e.[\mu]} g$ *iff* $\{x \in \mathrm{dom}(g) | f(x) = g(x)\}$ *is μ-conegligible*

If we have a sequence $\{f_n: X \to \mathbb{R}\}_{n\in\mathbb{N}}$ *of partial functions with* $\forall n\in\mathbb{N}\ \mathrm{dom}(f_n)\subseteq X$ *then we have*

1. $\lim_{n-<\infty} f_n =_{a.e.[\mu]} f$ *iff* $\left\{x\in \mathrm{dom}(f)\bigcap \mathrm{dom}\left(\lim_{n\to\infty} f_n\right)\middle| \lim_{n\to\infty} f_n = f\left(x\right\}\right.$ *is* μ-*conegligible where* $\mathrm{dom}\left(\lim_{n-<\infty} f_n\right) = \{x\in\bigcup_{n\in\mathbb{N}}\left(\bigcap_{i\in\{1,\ldots,n\}}\mathrm{dom}(f_i)\right)|\exists k\in\mathbb{N}$ such that $\{f_n(x)\}_{n\in\{k,\ldots\infty\}}$ has *a* limit in $\mathbb{R}\}$

Using the above definition we have the following rules to combine **almost every** predicates.

Proposition 18.30. *Let* $\langle X,\mathcal{A},\mu\rangle$ *be a measure space then we have*

1. *If* $f: X\to\mathbb{R}$, $g: X\to\mathbb{R}$, $h: X\to\mathbb{R}$ *are partial functions then we have*
   a. *If* $f =_{a.e.[\mu]} g$ *and* $g =_{a.e.[\mu]} h$ *then* $f =_{a.e.[\mu]} h$
   b. *If* $f \leqslant_{a.e.[\mu]} g$ *and* $g \leqslant_{a.e.[\mu]} h$ *then* $f \leqslant_{a.e.[\mu]} h$
   c. *If* $f <_{a.e.[\mu]} g$ *and* $g \leqslant_{a.e.[\mu]} h$ *then* $f <_{a.e.[\mu]} h$
   d. *If* $f \leqslant_{a.e.[\mu]} g$ *and* $g <_{a.e.[\mu]} h$ *then* $f <_{a.e.[\mu]} h$
   e. *If* $f <_{a.e.[\mu]} g$ *and* $g <_{a.e.[\mu]} h$ *then* $f <_{a.e.[\mu]} h$
2. *Let* $\{f_n: X\to\mathbb{R}\}_{n\in\mathbb{N}}$ *be partial functions so that* $\forall n\in\mathbb{N}$ *we have* $f_n \leqslant_{a.e.[\mu]} f_{n+1}$ *then* $\forall n,m\in\mathbb{N}$ *with* $n<m$ *we have* $f_n \leqslant_{a.e.[\mu]} f_m$
3. *If* $f: X\to\mathbb{R}$, $g: X\to\mathbb{R}$ *are partial functions then we have*
   a. $f =_{a.e.[\mu]} g$ *if and only* $f \leqslant_{a.e.[\mu]} g$ *and* $g \leqslant_{a.e.[\mu]} f$
   b. *Let* $a\in\mathbb{R}$ *then*
      i. *If* $a \leqslant_{a.e.[\mu]} f$ *then* $(-f) \leqslant_{a.e.[\mu]} -a$
      ii. *If* $f \leqslant_{a.e.[\mu]} a$ *then* $-a \leqslant_{a.e.[\mu]} (-f)$
      iii. $a =_{a.e.[\mu]} f$ *if and only if* $a \leqslant_{a.e.[\mu]} f$ *and* $f \leqslant_{a.e.[\mu]} a$
      iv. *If* $a \leqslant_{a.e.[\mu]} f$ *and* $f \leqslant_{a.e.[\mu]} g$ *then* $a \leqslant_{a.e.[\mu]} g$
      v. *If* $a <_{a.e.[\mu]} f$ *and* $f \leqslant_{a,e.[\mu]} g$ *then* $a \leqslant_{a.e.[\mu]} g$
      vi. *If* $a \leqslant_{a.e.[\mu]} f$ *and* $f <_{a.e.[\mu]} g$ *then* $a <_{a.e.[\mu]} g$
      vii. *If* $a <_{a.e.[\mu]} f$ *and* $f <_{a.e.[\mu]} g$ *then* $a <_{a.e.[\mu]} g$

Proof.

1.

   a. From $f \leqslant_{a.e.[\mu]} g$ we have that $E_1 = \{x\in\mathrm{dom}(f)\bigcap\mathrm{dom}(g)|f(x) = g(x)\}$ is conegligible and from $g =_{a.e.[\mu]} h$ we have $E_2 = \{x\in\mathrm{dom}(g)\bigcap\mathrm{dom}(h)|g(x) = h(x)\}$ is conegligible. Using 18.27 we have that $E_1\bigcap E_2$ is conegligible. Now $\forall x\in E_1\bigcap E_2\subseteq\mathrm{dom}(f)\bigcap\mathrm{dom}(g)\bigcap\mathrm{dom}(h)$ we have $f(x) = g(x)\wedge g(x) = h(x)\Rightarrow f(x) = h(x)$ proving that $E_1\bigcap E_2\subseteq\{x\in\mathrm{dom}(f)\bigcap\mathrm{dom}(h)|f(x) = h(x)\}$. Finally using 18.25 we have that $\{x\in\mathrm{dom}(f)\bigcap\mathrm{dom}(h)|f(x) = h(x)\}$ is conegligible proving that $f =_{a.e.[\mu]} h$.

b. From $f \leqslant_{a.e.[\mu]} g$ we have that $E_1 = \{x \in \text{dom}(f) \bigcap \text{dom}(g) | f(x) \leqslant g(x)\}$ is conegligible and from $g \leqslant_{a.e.[\mu]} h$ we have $E_2 = \{x \in \text{dom}(g) \bigcap \text{dom}(h) | g(x) \leqslant h(x)\}$ is conegligible. Using 18.27 we have that $E_1 \bigcap E_2$ is conegligible. Now $\forall x \in E_1 \bigcap E_2 \subseteq \text{dom}(f) \bigcap \text{dom}(g) \bigcap \text{dom}(h)$ we have $f(x) \leqslant g(x) \wedge g(x) \leqslant h(x) \Rightarrow f(x) \leqslant h(x)$ proving that $E_1 \bigcap E_2 \subseteq \{x \in \text{dom}(f) \bigcap \text{dom}(h) | f(x) \leqslant h(x)\}$. Finally using 18.25 we have that $\{x \in \text{dom}(f) \bigcap \text{dom}(h) | f(x) \leqslant h(x)\}$ is conegligible proving that $f \leqslant_{a.e.[\mu]} h$.

c. From $f <_{a.e.[\mu]} g$ we have that $E_1 = \{x \in \text{dom}(f) \bigcap \text{dom}(g) | f(x) < g(x)\}$ is conegligible and from $g \leqslant_{a.e.[\mu]} h$ we have $E_2 = \{x \in \text{dom}(g) \bigcap \text{dom}(h) | g(x) \leqslant h(x)\}$ is conegligible. Using 18.27 we have that $E_1 \bigcap E_2$ is conegligible. Now $\forall x \in E_1 \bigcap E_2 \subseteq \text{dom}(f) \bigcap \text{dom}(g) \bigcap \text{dom}(h)$ we have $f(x) < g(x) \wedge g(x) \leqslant h(x) \Rightarrow f(x) < h(x)$ proving that $E_1 \bigcap E_2 \subseteq \{x \in \text{dom}(f) \bigcap \text{dom}(h) | f(x) < h(x)\}$. Finally using 18.25 we have that $\{x \in \text{dom}(f) \bigcap \text{dom}(h) | f(x) < h(x)\}$ is conegligible proving that $f <_{a.e.[\mu]} h$.

d. From $f \leqslant_{a.e.[\mu]} g$ we have that $E_1 = \{x \in \text{dom}(f) \bigcap \text{dom}(g) | f(x) \leqslant g(x)\}$ is conegligible and from $g <_{a.e.[\mu]} h$ we have $E_2 = \{x \in \text{dom}(g) \bigcap \text{dom}(h) | g(x) < h(x)\}$ is conegligible. Using 18.27 we have that $E_1 \bigcap E_2$ is conegligible. Now $\forall x \in E_1 \bigcap E_2 \subseteq \text{dom}(f) \bigcap \text{dom}(g) \bigcap \text{dom}(h)$ we have $f(x) \leqslant g(x) \wedge g(x) < h(x) \Rightarrow f(x) < h(x)$ proving that $E_1 \bigcap E_2 \subseteq \{x \in \text{dom}(f) \bigcap \text{dom}(h) | f(x) < h(x)\}$. Finally using 18.25 we have that $\{x \in \text{dom}(f) \bigcap \text{dom}(h) | f(x) < h(x)\}$ is conegligible proving that $f <_{a.e.[\mu]} h$.

e. From $f <_{a.e.[\mu]} g$ we have that $E_1 = \{x \in \text{dom}(f) \bigcap \text{dom}(g) | f(x) < g(x)\}$ is conegligible and from $g <_{a.e.[\mu]} h$ we have $E_2 = \{x \in \text{dom}(g) \bigcap \text{dom}(h) | g(x) < h(x)\}$ is conegligible. Using 18.27 we have that $E_1 \bigcap E_2$ is conegligible. Now $\forall x \in E_1 \bigcap E_2 \subseteq \text{dom}(f) \bigcap \text{dom}(g) \bigcap \text{dom}(h)$ we have $f(x) < g(x) \wedge g(x) < h(x) \Rightarrow f(x) < h(x)$ proving that $E_1 \bigcap E_2 \subseteq \{x \in \text{dom}(f) \bigcap \text{dom}(h) | f(x) \leqslant h(x)\}$. Finally using 18.25 we have that $\{x \in \text{dom}(f) \bigcap \text{dom}(h) | f(x) < h(x)\}$ is conegligible proving that $f <_{a.e.[\mu]} h$.

2. We prove this by induction so given $n \in \mathbb{N}$ take $\mathcal{S}_n = \{n \in \mathbb{N} | f_n \leqslant_{a.e.[\mu]} f_{n+m}\}$ then we have:

   $1 \in \mathcal{S}_n$. this follows from the assumption $f_n \leqslant f_{N+1}$

   $m \in \mathcal{S}_n \Rightarrow m+1 \in \mathcal{S}_n$. now as $m \in \mathcal{S}_n$ we have that $f_n \leqslant_{a.e.[\mu]} f_{n+m}$ then as $f_{n+m} \leqslant_{a.e.[\mu]} f_{(n+m)+1}$ we have by (1)(a) that $f_n \leqslant_{a.e.[\mu]} f_{(n+m)+1} = f_{n+(m+1)}$ proving that $m+1 \in \mathcal{S}_n$

   Induction proves then that $\forall n \in \mathbb{N}$ we have $S_n = \mathbb{N}$ hence if $n, m \in S_n$ with $n < m$ then $1 \leqslant k = m - n$ so that $f_n \leqslant_{a.e.[\mu]} f_{n+k} = f_{n+(m-n)} = f_m$

3.

a. If $f =_{a.e.[\mu]} g$ then $E = \{x \in \mathrm{dom}(f) \bigcap \mathrm{dom}(g) | f(x) = g(x)\}$ is conegligible then $\forall x \in E$ we have $f(x) \leqslant g(x)$ and $g(x) \leqslant f(x)$ so that $E \subseteq \{x \in \mathrm{dom}(f) \bigcap \mathrm{dom}(g) | f(x) \leqslant g(x)\}, \{x \in \mathrm{dom}(f) \bigcap \mathrm{dom}(g) | g(x) \leqslant f(x)\}$ which by 18.25 proves that $\{x \in \mathrm{dom}(f) \bigcap \mathrm{dom}(g) | f(x) \leqslant g(x)\}$,$\{x \in \mathrm{dom}(f) \bigcap \mathrm{dom}(g) | g(x) \leqslant f(x)\}$ are conegligible. So $f \leqslant_{a.e.[\mu]} g$ and $g \leqslant_{a.e.[\mu]} f$.

For the opposite implication let $f \leqslant_{a.e.[\mu]} g$ and $g \leqslant_{a.e.[\mu]} f$ then $E_1 = \{x \in \mathrm{dom}(f) \bigcap \mathrm{dom}(g) | f(x) \leqslant g(x)\}$ is conegligible and $E_2 = \{x \in \mathrm{dom}(f) \bigcap \mathrm{dom}(g) | g(x) \leqslant f(x)\}$ is conegligible. Now $\forall x \in E_1 \bigcap E_2$ a conegligible set by 18.27 we have $f(x) \leqslant g(x)$ and $g(x) \leqslant f(x)$ giving $f(x) = g(x)$. This proves that $E_1 \bigcap E_2 \subseteq \{x \in \mathrm{dom}(f) \bigcap \mathrm{dom}(g) | f(x) = g(x)\}$ which by 18.27 proves that $\{x \in \mathrm{dom}(f) \bigcap \mathrm{dom}(g) | f(x) = g(x)\}$ is conegligible so that $f =_{a.e.[\mu]} g$

b.

i. As $a \leqslant_{a.e.[\mu]} f$ we have that $E = \{x \in \mathrm{dom}(f) | a \leqslant f(x)\}$ is conegligible. Now as

$$\begin{aligned} x \in E &\Leftrightarrow x \in \mathrm{dom}(f) \wedge a \leqslant f(x) \\ &\Leftrightarrow x \in \mathrm{dom}(f) \wedge -f(x) \leqslant -a \\ &\Leftrightarrow x \in \{\mathrm{dom}(f) | -f(x) \leqslant -a\} \end{aligned}$$

proving that $\{\mathrm{dom}(f) | -f(x) \leqslant -a\}$ is μ-conegligible and thus that $-f \leqslant_{a.e.[\mu]} -a$

ii. If $f \leqslant_{a.e.[\mu]} a$ then $-(-f) \leqslant_{a.e.[\mu]} -(-a)$ so by (3.b.i) we have $-a \leqslant_{a.e.[\mu]} -f$

iii. As $a =_{a.e.[\mu]} f$ we have that $E = \{x \in \mathrm{dom}(f) | a = f(x)\}$ is conegligible. Now $\forall x \in E$ we have $a \leqslant f(x)$ and $f(x) \leqslant a$ proving that $E \subseteq \{x \in \mathrm{dom}(f) | a \leqslant f(x)\}, \{x \in \mathrm{dom}(f) | f(x) \leqslant a\}$. Using 18.27 we have then that $\{x \in \mathrm{dom}(f) | a \leqslant f(x)\}, \{x \in \mathrm{dom}(f) | f(x) \leqslant a\}$ are conegligible proving that $a \leqslant_{a.e.[\mu]} f$ and $f \leqslant_{a.e.[\mu]} a$.

For the opposite implication, from $a \leqslant_{a.e.[\mu]} f$ and $f \leqslant_{a.e.[\mu]} a$ we have that $E_1 = \{x \in \mathrm{dom}(f) | a \leqslant f(x)\}$ and $E_2 = \{x \in \mathrm{dom}(f) | f(x) \leqslant a\}$ are conegligible. Using 18.27 it follows that $E_1 \bigcap E_2$ is conegligible. Now $\forall x \in E_1 \bigcap E_2$ we have $a \leqslant f(x)$ and $f(x) \leqslant a \Rightarrow f(x) = a$ which proves that $E_1 \bigcap E_2 \subseteq \{x \in \mathrm{dom}(f) | a = f(x)\}$. Applying 18.27 proves that $\{x \in \mathrm{dom}(f) | a = f(x)\}$ is conegligible and thus that $a =_{a.e.[\mu]} f$.

iv. From $a \leqslant_{a.e.[\mu]} f$ and $f \leqslant_{a.e.[\mu]} g$ it follows that $E_1 = \{x \in \mathrm{dom}(f) | a \leqslant f(x)\}$ and $E_2 = \{x \in \mathrm{dom}(f) \bigcap \mathrm{dom}(g) | f(x) \leqslant g(x)\}$ are conegligible. By 18.27 we have then that $E_1 \bigcap E_2$ is conegligible. Now $\forall x \in E_1 \bigcap E_2$ we have that $a \leqslant f(x) \leqslant g(x)$ hence $a \leqslant g(x)$ which proves that $E_1 \bigcap E_2 \subseteq \{x \in \mathrm{dom}(g) | a \leqslant g(x)\}$. Applying then 18.27 proves that $\{x \in \mathrm{dom}(g) | a \leqslant g(x)\}$ and thus that $a \leqslant_{a.e.[\mu]} g$.

v. From $a <_{a.e.[\mu]} f$ and $f \leqslant_{a.e.[\mu]} g$ it follows that $E_1 = \{x \in \mathrm{dom}(f) | a < f(x)\}$ and $E_2 = \{x \in \mathrm{dom}(f) \bigcap \mathrm{dom}(g) | f(x) \leqslant g(x)\}$ are conegligible. By 18.27 we have then that $E_1 \bigcap E_2$ is conegligible. Now $\forall x \in E_1 \bigcap E_2$ we have that $a < f(x) \leqslant g(x)$ hence $a < g(x)$ which proves that $E_1 \bigcap E_2 \subseteq \{x \in \mathrm{dom}(g) | a < g(x)\}$. Applying then 18.27 proves that $\{x \in \mathrm{dom}(g) | a < g(x)\}$ and thus that $a <_{a.e.[\mu]} g$.

vi. From $a \leqslant_{a.e.[\mu]} f$ and $f <_{a.e.[\mu]} g$ it follows that $E_1 = \{x \in \mathrm{dom}(f) | a \leqslant f(x)\}$ and $E_2 = \{x \in \mathrm{dom}(f) \bigcap \mathrm{dom}(g) | f(x) < g(x)\}$ are conegligible. By 18.27 we have then that $E_1 \bigcap E_2$ is conegligible. Now $\forall x \in E_1 \bigcap E_2$ we have that $a \leqslant f(x) < g(x)$ hence $a < g(x)$ which proves that $E_1 \bigcap E_2 \subseteq \{x \in \mathrm{dom}(g) | a < g(x)\}$. Applying then 18.27 proves that $\{x \in \mathrm{dom}(g) | a \leqslant g(x)\}$ and thus that $a <_{a.e.[\mu]} g$.

vii. From $a <_{a.e.[\mu]} f$ and $f <_{a.e.[\mu]} g$ it follows that $E_1 = \{x \in \mathrm{dom}(f) | a < f(x)\}$ and $E_2 = \{x \in \mathrm{dom}(f) \bigcap \mathrm{dom}(g) | f(x) < g(x)\}$ are conegligible. By 18.27 we have then that $E_1 \bigcap E_2$ is conegligible. Now $\forall x \in E_1 \bigcap E_2$ we have that $a < f(x) < g(x)$ hence $a < g(x)$ which proves that $E_1 \bigcap E_2 \subseteq \{x \in \mathrm{dom}(g) | a < g(x)\}$. Applying then 18.27 proves that $\{x \in \mathrm{dom}(g) | a < g(x)\}$ and thus that $a <_{a.e.[\mu]} g$.

□

18.1.4 Carathéodory construction

The following lemma will be useful in some of the proofs that follow this, it allows us to avoid having to use a proof by induction argument in these proofs.

Lemma 18.31. *Let X be a set and $\mathcal{A} \subseteq \mathcal{P}(X)$ then we have*

1. *If $\forall A, B \in \mathcal{A}$ we have $A \bigcup B$ then $\forall \{A_i\}_{i \in \{1,\ldots,n\}}$ where $n \in \mathbb{N}$ we have $\bigcup_{i \in \{1,\ldots,n\}} A_i \in \mathcal{A}$*
2. *If $\forall A, B \in \mathcal{A}$ we have $A \bigcup B$ then $\forall \{A_i\}_{i \in \{1,\ldots,n\}}$ where $n \in \mathbb{N}$ we have $\bigcup_{i \in \{1,\ldots,n\}} A_i \in \mathcal{A}$*

Proof. We use a mathematical induction in this proof.

1. Let $\mathcal{S} = \left\{n \in \mathbb{N} | \forall \{A_i\}_{i \in \{1,\ldots,n\}} \subseteq \mathcal{A} \text{ we have } \bigcup_{i \in \{1,\ldots,n\}} A_i\right\}$ then we have

   $1 \in \mathcal{S}$. if $\{A_i\}_{i \in \{1,\ldots,1\}}$ then $\bigcup_{i \in \{1,\ldots,1\}} A_i = A_1 \in \mathcal{A}$

   $n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$. if $\{A_i\}_{i \in \{1,\ldots,n+1\}} \subseteq \mathcal{A}$ then as $n \in \mathcal{S}$ we have that $\bigcup_{i \in \{1,\ldots,n\}} A_i \in \mathcal{A}$ hence $\bigcup_{i \in \{1,\ldots,n+1\}} A_i = \left(\bigcup_{i \in \{1,\ldots,n\}} A_i\right) \bigcup A_{n+1} \in \mathcal{A}$

2. Let $\mathcal{S} = \left\{n \in \mathbb{N} | \forall \{A_i\}_{i \in \{1,\ldots,n\}} \subseteq \mathcal{A} \text{ we have } \bigcap_{i \in \{1,\ldots,n\}} A_i\right\}$ then we have

   $1 \in \mathcal{S}$. if $\{A_i\}_{i \in \{1,\ldots,1\}}$ then $\bigcap_{i \in \{1,\ldots,1\}} A_i = A_1 \in \mathcal{A}$

   $n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$. if $\{A_i\}_{i \in \{1,\ldots,n+1\}} \subseteq \mathcal{A}$ then as $n \in \mathcal{S}$ we have that $\bigcap_{i \in \{1,\ldots,n\}} A_i \in \mathcal{A}$ hence $\bigcap_{i \in \{1,\ldots,n+1\}} A_i = \left(\bigcap_{i \in \{1,\ldots,n\}} A_i\right) \bigcup A_{n+1} \in \mathcal{A}$ □

Constructing measures on a σ-algebra is not a easy task. The idea of the Carathéodory construction is that is that we first start with a set function that has some of the properties that a measure must have (see 18.20) and then find a σ-algebra that makes this a measure on this σ-algebra. With luck this σ-algebra is the σ-algebra where we are constructing a measure for.

First we define the concept of outer measure that is a set function that fulfills the minimum requirements that a measure must have.

Definition 18.32. (Outer Measure) *Let X be a set then $\mathcal{O}:\mathcal{P}(X)\rightarrow[0,\infty]$ is a* ***outer*** *measure on X iff*

1. $\mathcal{O}(\emptyset)=0$
2. *If $A,B\in\mathcal{P}(X)$ with $A\subseteq B$ then $\mathcal{O}(A)\leqslant\mathcal{O}(B)$*
3. *$\forall\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{P}(X)$ then $\mathcal{O}(\bigcup_{i\in\mathbb{N}}A_i)\leqslant\sum_{i\in\mathbb{N}}\mathcal{O}(A_i)$*

The following shows how to construct a outer measure based on a semi-additive function

Theorem 18.33. *Let X be a set, $\mathcal{A}\subseteq\mathcal{P}(X)$ and $\mu:\mathcal{A}\rightarrow[0,\infty]$ such that*

1. $\emptyset\in\mathcal{A}$
2. *$\exists\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{A}$ such that $X\subseteq\bigcup_{i\in\mathbb{N}}A_i$*
3. $\mu(\emptyset)=0$
4. *$\forall A\in\mathcal{A}$, $\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{A}$ we have $\mu(A)\leqslant\sum_{i=1}^{\infty}\mu(A_i)$*

then $\varphi:\mathcal{P}(X)\rightarrow[0,\infty]$ defined by $\varphi(A)=\inf(\{\sum_{i=1}^{\infty}\mu(A_i)|\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{A}\,\text{such that}\,A\subseteq\bigcup_{i\in\mathbb{N}}A_i\})$ is a outer measure on X. Further $\forall A\in\mathcal{A}$ we have $\varphi(A)=\mu(A)$

Proof. First we must prove that φ is defined for every $A\in\mathcal{P}(X)$. Let

$$\mathcal{C}_A=\left\{\sum_{i=1}^{\infty}\mu(A_i)|\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{A}\,\text{such that}\,A\subseteq\bigcup_{i\in\mathbb{N}}A_i\right\}\quad\text{where}\,A\in\mathcal{P}(X)\tag{18.20}$$

Given $A\in\mathcal{P}(X)$ there exists using (2) a $\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{A}$ such that $A\subseteq X\subseteq\bigcup_{i\in\mathbb{N}}A_i$ proving that $\mathcal{C}_A\neq\emptyset$ so by 17.10 $\lim(C_A)$ exists and as 0 is a a lower bound of $\mathcal{C}_A$ we have that

$$\forall A\in\mathcal{P}(X)\;\varphi(A)=\inf(\mathcal{C}_A)\,\text{exists}\;\text{and}\;\varphi(A)\in[0,\infty]\tag{18.21}$$

Next we check the conditions for a outer measure.

1. Define $\{A_i\}_{i\in\mathbb{N}}$ by $A_i=\emptyset$ then by (1) $\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{A}$ and $\emptyset\subseteq\bigcup_{i\in\mathbb{N}}A_i$. As $\sum_{i=1}^{\infty}\mu(A_i)\underset{(3)}{=}\sum_{i=1}^{\infty}0=0$ we have that $0\leqslant\varphi(\emptyset)\leqslant0$. Hence
$$\varphi(\emptyset)\tag{18.22}$$
2. Let $A,B\in\mathcal{P}(X)$ with $A\subseteq B$ then if $x\in\mathcal{C}_B$ there exists a $\{B_i\}_{i\in\mathbb{N}}\subseteq\mathcal{A}$ such that $B\subseteq\bigcup_{i\in\mathbb{N}}B_i$ and $x=\sum_{i=1}^{\infty}\mu(B_i)$. As $A\subseteq B\subseteq\bigcup_{i\in\mathbb{N}}B_i$ it follows that $x\in\mathcal{C}_A$ or $\mathcal{C}_B\subseteq\mathcal{C}_A$. So using 2.171 we have that $\inf(\mathcal{C}_A)\leqslant\inf(\mathcal{C}_B)$ or
$$\varphi(A)\leqslant\varphi(B)\,\text{for}\,A\subseteq B\tag{18.23}$$

3. Let $\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{P}(X)$ and note $A=\bigcup_{i\in\mathbb{N}}A_i$. Take $\varepsilon>0$ then using the definition of φ and the infinum there exists a $\forall n\in\mathbb{N}$ a sequence $\{I_{n,i}\}_{n,i\in\mathbb{N}}\subseteq\mathcal{A}$ such that

$$\varphi(A_n)\leqslant\sum_{i=1}^{\infty}\mu(I_{n,i})<\varphi(A_n)+\frac{\varepsilon}{2^n}\text{ and }A_n\subseteq\bigcup_{i\in\mathbb{N}}I_{n,i}\tag{18.24}$$

As $\mathbb{N}\times\mathbb{N}$ is denumerable (or infinite countable) (see 5.59) there exists a bijection $\beta:\mathbb{N}\to\mathbb{N}\times\mathbb{N}$. We prove now that

$$A\subseteq\bigcup_{n\in\mathbb{N}}I_{\beta(n)_1,\beta(n)_2}\tag{18.25}$$

Proof. If $x\in A$ then $\exists n\in\mathbb{N}$ such that $x\in A_n\subseteq\bigcup_{i\in\mathbb{N}}I_{n,i}$, hence $\exists i\in\mathbb{N}$ such that $x\in I_{n,i}$. As φ is surjective there exists a $k\in\mathbb{N}$ such that $\beta(k)=(n,i)$ so that $x\in I_{\beta(k)_1,\beta(k)_2}\subseteq\bigcup_{k\in\mathbb{N}}I_{\beta(k)_1,\beta(k)_2}$ □

Now

$$\begin{aligned}
\varphi(A) &= \inf(\mathcal{C}_A)\\
&\underset{(18.20)}{=} \inf\left(\left\{\sum_{i=1}^{\infty}\mu(A_i)|\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{A}\,\text{such that}\,A\subseteq\bigcup_{i\in\mathbb{N}}A_i\right\}\right)\\
&\leqslant_{(18.25)} \sum_{i=1}^{\infty}\mu(I_{\beta(n)_1,\beta(n)_2})\\
&\underset{17.124}{=} \sum_{i=1}^{\infty}\left(\sum_{j=1}^{\infty}\mu(I_{i,j})\right)\\
&\leqslant_{(18.24)} \sum_{i=1}^{\infty}\left(\varphi(A_i)+\frac{\varepsilon}{2^i}\right)\\
&\underset{17.114}{=} \sum_{i=1}^{\infty}\varphi(A_i)+\sum_{i=1}^{\infty}\frac{\varepsilon}{2^i}\\
&= \sum_{i=1}^{\infty}\varphi(A_i)+\varepsilon\cdot\sum_{i=1}^{\infty}\frac{1}{2^i}\\
&\underset{12.403}{=} \sum_{i=1}^{\infty}\varphi(A_i)+\varepsilon\cdot\frac{\frac{1}{2}}{1-\frac{1}{2}}\\
&= \sum_{i=1}^{\infty}\varphi(A_i)+\varepsilon
\end{aligned}$$

As ε is chosen arbitrary we have using 9.56 that

$$\varphi\left(\bigcup_{i\in\mathbb{N}}A_i\right)=\varphi(A)\leqslant\sum_{i=1}^{\infty}\varphi(A_i)\tag{18.26}$$

Using (18.22),(18.23) and the above it follows that φ is a outer measure. For the last part of the theorem let $A\in\mathcal{A}$ and define $\{A_i\}_{i\in\mathbb{N}}$ by $A_i=\left\{\begin{smallmatrix} A \text{ if } i=1 \\ \emptyset \text{ if } i\in\mathbb{N}\setminus\{1\}\end{smallmatrix}\right.\in\mathcal{A}$ then clearly $A\subseteq\bigcup_{i\in\mathbb{N}}A_i$ so that $\varphi(A)=\sup(\mathcal{C}_A)\leqslant\sum_{i=1}^{\infty}\mu(A_i)=\mu(A_1)=\mu(A)$. Further if $\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{P}(X)$ is such that $A\subseteq\bigcup_{i\in\mathbb{N}}A_i$ then using (4) we have that $\mu(A)\leqslant\sum_{i\in\mathbb{N}}\mu(A_i)$ so that $\mu(A)\leqslant\inf(\mathcal{C}_A)=\varphi(A)$. Hence we have

$$\forall A\in\mathcal{A}\text{ we have }\varphi(A)=\mu(A)$$

□

We are now ready to construct given a set X, a outer measure on X a σ-algebra on X such that the restriction of the outer measure on this σ-algebra is a measure.

Theorem 18.34. (Carathéodory) *Let X be a set and $\mathcal{O}:\mathcal{P}(X)\rightarrow[0,\infty]$ a outer measure on X then $\mathcal{A}=\{E\in\mathcal{P}(X)|\forall A\in\mathcal{P}(X)\text{ we have }\mathcal{O}(A)=\mathcal{O}(A\bigcap E)+\mathcal{O}(A\setminus E)\}$ is a σ-algebra and $\mathcal{O}_{|\mathcal{A}}:\mathcal{A}\rightarrow[0,\infty]$ is a measure, making $\langle X,\mathcal{A},\mathcal{O}_{|\mathcal{A}}\rangle$ a measure space. Further $\mathcal{A}=\{E\in\mathcal{P}(X)|\forall A\in\mathcal{P}(X)\text{ we have }\mathcal{O}(A)\geqslant\mathcal{O}(A\bigcap E)+\mathcal{O}(A\setminus E)\}$*

Proof. First we prove that

$$\mathcal{A}=\left\{E\in\mathcal{P}(X)|\forall A\in\mathcal{P}(X)\text{ we have }\mathcal{O}(A)\geqslant\mathcal{O}\left(A\bigcap E\right)+\mathcal{O}(A\setminus E)\right. \tag{18.27}$$

Proof. First if $E\in\mathcal{A}$ then $\forall A$ we have $\mathcal{O}(A)=\mathcal{O}(A\bigcap E)+\mathcal{O}(A\setminus E)$ hence $\mathcal{O}(A)\geqslant\mathcal{O}(A\bigcap E)+\mathcal{O}(A\setminus E)$ so that $E\in\{E\in\mathcal{P}(X)|\forall A\in\mathcal{P}(X)\text{ we have }\mathcal{O}(A)\geqslant\mathcal{O}(A\bigcap E)+\mathcal{O}(A\setminus E)\}$, second if $E\in\{E\in\mathcal{P}(X)|\forall A\in\mathcal{P}(X)\text{ we have }\mathcal{O}(A)\geqslant\mathcal{O}(A\bigcap E)+\mathcal{O}(A\setminus E)\}$ then $\forall A$ we have $\mathcal{O}(A)\geqslant\mathcal{O}(A\bigcap E)+\mathcal{O}(A\setminus E)$. As $A=(A\bigcap E)\bigcup(A\setminus E)$ it follows from 18.32 (2) that $\mathcal{O}(A)\leqslant\mathcal{O}(A\bigcap E)+\mathcal{O}(A\setminus E)$ and thus that $\mathcal{O}(A)=\mathcal{O}(A\bigcap E)+\mathcal{O}(A\setminus E)$. Hence $E\in\mathcal{A}$. □

Next we prove that $\mathcal{A}$ is a σ-algebra.

1. Let $A\in\mathcal{P}(X)$ then $A\bigcap\emptyset=\emptyset$ and $A\setminus\emptyset=A$ so that $\mathcal{O}(A)=0+\mathcal{O}(A\setminus\emptyset)\underset{18.32\ (1)}{=}\mathcal{O}(\emptyset)+\mathcal{O}(A\setminus\emptyset)$ proving that

$$\emptyset\in\mathcal{A} \tag{18.28}$$

2. Let $E\in\mathcal{A}$ then $\forall A\in\mathcal{P}(X)$ we have

$$\begin{aligned}
\mathcal{O}\left(A\bigcap(X\setminus E)\right)+\mathcal{O}(A\setminus(X\setminus E)) &= \mathcal{O}\left(A\bigcap\left(X\bigcap E^c\right)\right)+\mathcal{O}(A\setminus(X\setminus E)) \\
&= \mathcal{O}\left(A\bigcap E^c\right)+\mathcal{O}(A\setminus(X\setminus E)) \\
&= \mathcal{O}(A\setminus E)+\mathcal{O}(A\setminus(X\setminus E)) \\
&= \mathcal{O}(A\setminus E)+\mathcal{O}\left(A\bigcap(X\setminus E)^c\right) \\
&= \mathcal{O}(A\setminus E)+\mathcal{O}\left(A\bigcap\left(X\bigcap E^c\right)^c\right) \\
&= \mathcal{O}(A\setminus E)+\mathcal{O}\left(A\bigcap\left(X^c\bigcup(E^c)^c\right)\right) \\
&= \mathcal{O}(A\setminus E)+\mathcal{O}\left(A\bigcap\left(X^c\bigcup E\right)\right) \\
&= \mathcal{O}(A\setminus E)+\mathcal{O}\left(\left(A\bigcap X^c\right)\bigcup\left(A\bigcap E\right)\right) \\
&\underset{A\subseteq X}{=} \mathcal{O}(A\setminus E)+\mathcal{O}\left(A\bigcap E\right) \\
&\underset{E\in\mathcal{A}}{=} \mathcal{O}(A)
\end{aligned}$$

proving that

$$\forall E \in \mathcal{A} \text{ we have } X \setminus E \in \mathcal{A} \tag{18.29}$$

3. Let $E, F \in \mathcal{A}$ then

$$\begin{aligned}
\mathcal{O}\left(A\bigcap\left(E\bigcup F\right)\right) &\underset{E\in\mathcal{A},A\bigcap(E\bigcup F)\in\mathcal{P}(X)}{=} \mathcal{O}\left(\left(A\bigcap\left(E\bigcup F\right)\right)\bigcap E\right)+\mathcal{O}\left(\left(A\bigcap\left(E\bigcup F\right)\right)\setminus E\right)\\
&\underset{1.31}{=} \mathcal{O}\left(A\bigcap E\right)+\mathcal{O}\left(\left(A\bigcap\left(E\bigcup F\right)\right)\setminus E\right)\\
&\underset{1.31}{=} \mathcal{O}\left(A\bigcap E\right)+\mathcal{O}\left(A\bigcap\left((E\setminus E)\bigcup(F\setminus E)\right)\right)\\
&= \mathcal{O}\left(A\bigcap E\right)+\mathcal{O}\left(A\bigcap(F\setminus E)\right)\\
&\underset{1.31}{=} \mathcal{O}\left(A\bigcap E\right)+\mathcal{O}\left((A\setminus E)\bigcap F\right)
\end{aligned} \tag{18.30}$$

so

$$\begin{aligned}
\mathcal{O}\left(A\bigcap\left(E\bigcup F\right)\right)+\mathcal{O}\left(A\setminus\left(E\bigcup F\right)\right) &\underset{18.30}{=} \mathcal{O}\left(A\bigcap E\right)+\mathcal{O}\left((A\setminus E)\bigcap F\right)+\mathcal{O}\left(A\setminus\left(E\bigcup F\right)\right)\\
&\underset{1.31}{=} \mathcal{O}\left(A\bigcap E\right)+\mathcal{O}\left((A\setminus E)\bigcap F\right)+\mathcal{O}((A\setminus E)\setminus F)\\
&\underset{(A\setminus F)\in\mathcal{P}(X)\wedge F\in\mathcal{A}}{=} \mathcal{O}\left(A\bigcap E\right)+\mathcal{O}(A\setminus E)\\
&\underset{E\in\mathcal{A}}{=} \mathcal{O}(A)
\end{aligned}$$

proving that

$$\forall E, F \in \mathcal{A} \text{ we have } E\bigcup F \in \mathcal{A} \tag{18.31}$$

4. Let $\{E_i\}_{i\in\mathbb{N}} \subseteq \mathcal{A}$. Define then $\{G_n\}_{n\in\mathbb{N}}$ by $G_n = \bigcup_{i\in\{1,\ldots,n\}} E_i$ and $\{F_n\}_{n\in\mathbb{N}}$ by $F_i = \left\{\begin{smallmatrix} G_1 \text{ if } i=1 \\ G_i\setminus G_{i-1} \text{ if } i\in\mathbb{N}\setminus\{1\}\end{smallmatrix}\right.$. By construction we have

$$\forall n \in \mathbb{N} \text{ that } G_n \subseteq G_{n+1} \tag{18.32}$$

Further

$$\bigcup_{i\in\mathbb{N}} E_i = \bigcup_{i\in\mathbb{N}} G_i = \bigcup_{i\in\mathbb{N}} F_i \tag{18.33}$$

Proof. As $\forall n \in \mathbb{N}$ we have $G_n = \bigcup_{i\in\{1,\ldots,n\}} A_i \subseteq \bigcup_{i\in\mathbb{N}} E_i$ we have that $\bigcup_{i\in\mathbb{N}} G_i \subseteq \bigcup_{i\in\mathbb{N}} E_i$. Further if $x \in \bigcup_{i\in\mathbb{N}} E_i$ then there exists a $i\in\mathbb{N}$ such that $x \in E_i \subseteq \bigcup_{j\in\{1,\ldots,i\}} E_j = G_i$ so that $\bigcup_{i\in\mathbb{N}} E_i \subseteq \bigcup_{i\in\mathbb{N}} G_i$. Next as $F_1 = G_1$ and $\forall i \in \mathbb{N}\setminus\{1\}$ we have $F_i = G_i\setminus G_{i-1} \subseteq G_i$ we have that $\bigcup_{i\in\mathbb{N}} F_i \subseteq \bigcup_{i\in\mathbb{N}} G_i$. Finally if $x \in \bigcup_{i\in\mathbb{N}} G_i$ then $\exists i \in \mathbb{N}$ such that $x \in G_i$ and then either $i = 1$ so that $x \in F_1$ or $1 < i$ giving $x \in F_i = G_i \setminus G_{i-1} \subseteq G_i$ proving that $\bigcup_{i\in\mathbb{N}} G_i \subseteq \bigcup_{i\in\mathbb{N}} F_i$. $\square$

Using 18.31 and 18.31 we have

$$\{G_i\}_{i\in\mathbb{N}}\subseteq\mathcal{A} \tag{18.34}$$

Take $A\in\mathcal{P}(X)$, $n\in\mathbb{N}\setminus\{1\}$ then as $G_{n-1}\in\mathcal{A}$ and $A\bigcap G\in\mathcal{P}(X)$ we have

$$\begin{aligned}\mathcal{O}(A\bigcap G_n) &= \mathcal{O}((A\bigcap G_n)\bigcap G_{n-1})+\mathcal{O}((A\bigcap G_n)\setminus G_{n-1})\\ &\underset{1.31}{=} \mathcal{O}(A\bigcap (G_n\bigcap G_{n-1}))+\mathcal{O}(A\bigcap (G_n\setminus G_{n-1}))\\ &\underset{18.33}{=} \mathcal{O}(A\bigcap G_{n-1})+\mathcal{O}(A\bigcap F_n)\end{aligned} \tag{18.35}$$

We prove now by induction that

$$\forall n\in\mathbb{N},\forall A\in\mathcal{P}(X)\vDash\mathcal{O}(A\bigcap G_n)=\sum_{i=1}^{n}\mathcal{O}(A\bigcap F_i) \tag{18.36}$$

Proof. Let $\mathcal{S}=\left\{n\in\mathbb{N}|\mathcal{O}(A\bigcap G_n)=\sum_{i\in\{1,\ldots,n\}}\mathcal{O}(A\bigcap F_i)\right\}$ then we have

$1\in\mathcal{S}$. $\mathcal{O}[A\bigcap G_1]\underset{\text{def}}{=}\mathcal{O}(A\bigcap F_1)=\sum_{i=1}^{1}\mathcal{O}(A\bigcap F_i)$ proving that $1\in\mathcal{S}$

$n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$. then

$$\begin{aligned}\mathcal{O}[A\bigcap G_{n+1}] &\underset{18.35}{=} \mathcal{O}(A\bigcap G_n)+\mathcal{O}(A\bigcap F_{n+1})\\ &\underset{n\in\mathcal{S}}{=} \sum_{i=1}^{n}\mathcal{O}(A\bigcap F_i)+\mathcal{O}(A\bigcap F_{n+1})\\ &= \sum_{i=1}^{n+1}\mathcal{O}(A\bigcap F_i)\end{aligned}$$

proving $n+1\in\mathcal{S}$ □

Let $A\in\mathcal{P}(X)$ then $A\bigcap\left(\bigcup_{i\in\mathbb{N}}E_i\right)\underset{18.33}{=}A\bigcap\left(\bigcup_{i\in\mathbb{N}}F_i\right)=\bigcup_{i\in\mathbb{N}}(A\bigcap F_i)$ so that as $\mathcal{O}$ is a outer measure we have

$$\begin{aligned}\mathcal{O}\left(A\bigcap\left(\bigcup_{i\in\mathbb{N}}E_i\right)\right) &\leqslant \sum_{i=1}^{\infty}\mathcal{O}(A\bigcap F_i)\\ &= \lim_{i\to\infty}\left(\sum_{i=1}^{n}\mathcal{O}(A\bigcap F_i)\right)\\ &\underset{18.36}{=} \lim_{i\to\infty}\mathcal{O}(A\bigcap G_i)\end{aligned} \tag{18.37}$$

Next as $\forall n\in\mathbb{N}$ we have that

$$\begin{aligned}A\setminus\left(\bigcup_{i\in\mathbb{N}}E_i\right) &\underset{18.33}{=} A\setminus\left(\bigcup_{i\in\mathbb{N}}G_i\right)\\ &\underset{1.108}{=} \bigcap_{i\in\mathbb{N}}(A\setminus G_i)\\ &\subseteq A\setminus G_n\end{aligned}$$

and as $\mathcal{O}$ is a outer measure we have that $\mathcal{O}(A\setminus(\bigcup_{i\in\mathbb{N}} E_i)) \subseteq \mathcal{O}(A\setminus G_n)$ proving that

$$\mathcal{O}\left(A\setminus\left(\bigcup_{i\in\mathbb{N}} E_i\right)\right) \leqslant \inf\left(\{\mathcal{O}(A\setminus G_n)|n\in\mathbb{N}\}\right) \tag{18.38}$$

Now given $n\in\mathbb{N}$ we have using 18.32 that $G_n\subseteq G_{n+1}$ hence $X\setminus G_{n+1}\subseteq X\setminus G_n$, using the fact that $\mathcal{O}$ is a outer measure we must have that $\mathcal{O}(X\setminus G_{n+1})\leqslant\mathcal{O}(X\setminus G_n)$. So $\{\mathcal{O}(X\setminus G_i)\}_{i\in\mathbb{N}}$ is decreasing and thus by 17.83 we have that $\inf\Big(\{\mathcal{O}(A\setminus G_n)|n\in\mathbb{N}\} = \lim_{i\to\infty}\mathcal{O}(X\setminus G_i)$. Using 18.38 we have then that

$$\mathcal{O}\left(A\setminus\left(\bigcup_{i\in\mathbb{N}} E_i\right)\right) \leqslant \lim_{i\to\infty}\mathcal{O}(X\setminus G_i) \tag{18.39}$$

Applying now 18.37 and 18.39 gives

$$\begin{aligned}
\mathcal{O}\left(A\bigcap\left(\bigcup_{i\in\mathbb{N}} E_i\right)\right)+\mathcal{O}\left(A\setminus\left(\bigcup_{i\in\mathbb{N}} E_i\right)\right) &\leqslant \lim_{i\to\infty}\mathcal{O}\big(A\bigcap G_i\big)+\lim_{i\to\infty}\mathcal{O}(X\setminus G_i)\\
&\underset{17.89}{=} \lim_{i\to\infty}\big(\mathcal{O}\big(A\bigcap G_i\big)+\mathcal{O}(X\setminus G_i)\big)\\
&\underset{G_i\in\mathcal{A}}{=} \lim_{i\to}\mathcal{O}(A)\\
&= \mathcal{O}(A)
\end{aligned} \tag{18.40}$$

Using this with 18.27 we have that $\bigcup_{i\in\mathbb{N}} E_i\in\mathcal{A}$ proving that

$$\forall\{E_i\}_{i\in\mathbb{N}}\subseteq\mathcal{A}\,\text{we have that}\bigcup_{i\in\mathbb{N}} E_i\in\mathcal{A} \tag{18.41}$$

Combining 18.28, 18.29 and 18.41 we have proved that

$$\mathcal{A}=\big\{E\in\mathcal{P}(X)|\forall A\in\mathcal{P}(X)\text{ we have }\mathcal{O}(A)\geqslant\mathcal{O}\big(A\bigcap E\big)+\mathcal{O}(A\setminus E)\text{ is a }\sigma\text{-algebra} \tag{18.42}$$

Finally we must prove that the restriction of $\mathcal{O}$ to $\mathcal{A}$ is a measure. Define μ: $\mathcal{A}\to[0,\infty]$ by $\mu=\mathcal{O}_{\mathcal{A}}$. Then we have that $\mu(\emptyset)=\mathcal{O}_{|\mathcal{A}}(\emptyset)=\mathcal{O}(\emptyset=0$ proving that

$$\mu(\emptyset)=0 \tag{18.43}$$

Let now $\{E_i\}_{i\in\mathbb{N}}\subseteq\mathcal{A}$ be such that $\forall i,j\in\mathbb{N}$ with $i\neq j$ $A_i\bigcap A_j=\emptyset$. If we define then $\{G_i\}_{i\in\mathbb{N}}$ by $G_i=\bigsqcup_{j\in\{1,\ldots,i\}} E_i$ then using (6) 18.32,18.33 we have

$$\bigsqcup_{i\in\mathbb{N}} E_i=\bigcup_{i\in\mathbb{N}} G_i\ \text{ and }\ \forall i\in\mathbb{N}\,\text{we have}\,G_i\subseteq G_{i+1} \tag{18.44}$$

but also as $\mathcal{A}$ is a σ-algebra that

$$\{G_i\}_{i\in\mathbb{N}}\subseteq\mathcal{A}\,\text{and}\bigcup_{i\in\mathbb{N}} E_i\in\mathcal{A} \tag{18.45}$$

Given $n \in \mathbb{N}$ we have that

$$\begin{aligned}
G_{n+1} \setminus E_{n+1} &= \left(\left(\bigcup_{i \in \{1,\ldots,n+1\}} E_i\right) \setminus E_{n+1}\right) \\
&= \left(\left(\bigcup_{i \in \{1,\ldots,n\}} E_i\right) \bigcup E_{n+1}\right) \setminus E_{n+1} \\
&\underset{1.108}{=} \left(\left(\bigcup_{i \in \{1,\ldots,n\}} E_i\right) \setminus E_{n+1}\right) \bigcup (E_{n+1} \setminus E_{n+1}) \\
&\underset{1.108}{=} \bigcup_{i \in \{1,\ldots,n\}} (E_i \setminus E_{n+1}) \\
&\underset{\forall i \in \{1,\ldots,n\} A_{n+1} \bigcap A_i = \emptyset i}{=} \bigcup_{i \in \{1,\ldots,n\}} E_i \quad \text{(see 1.32))} \\
&= G_n
\end{aligned} \tag{18.46}$$

so

$$\begin{aligned}
\mu(G_{n+1}) &= \mathcal{O}(G_{n+1}) \\
&\underset{E_{n+1} \in \mathcal{A} \wedge G_{n+1} \in \mathcal{P}(X)}{=} \mathcal{O}\left(G_{n+1} \bigcap E_{n+1}\right) + \mathcal{O}(G_{n+1} \setminus E_{n+1}) \\
&\subseteq_{E_{n+1} \subseteq G_{n+1}} \mathcal{O}(E_{n+1}) + \mathcal{O}(G_{n+1} \setminus E_{n+1}) \\
&\underset{18.46}{=} \mathcal{O}(E_{n+1}) + \mathcal{O}(G_n) \\
& \mu(E_{n+1}) + \mu(G_n)
\end{aligned} \tag{18.47}$$

We prove now by induction that

$$\forall n \in \mathbb{N} \text{ we have } \mu\left(\bigsqcup_{i \in \{1,\ldots,n\}} E_i\right) = \sum_{i=1}^{n} \mu(E_i) \tag{18.48}$$

Proof. Define $\mathcal{S} = \left\{n \in \mathbb{N} | \mu\left(\bigsqcup_{i \in \{1,\ldots,n\}} E_i\right) = \sum_{i=1}^{n} \mu(A_i)\right\}$ then we have

$1 \in \mathcal{S}$. $\mu\left(\bigcup_{i \in \{1,\ldots,1\}} E_i\right) = \mu(E_1) = \sum_{i=1}^{1} \mu(E_i)$ proving that $1 \in \mathcal{S}$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$.

$$\begin{aligned}
\mu\left(\bigsqcup_{i \in \{1,\ldots,n+1\}} E_i\right) &= \mu(G_{n+1}) \\
&\underset{18.47}{=} \mu(E_{n+1}) + \mu(G_n) \\
&\underset{n \in \mathcal{S}}{=} \mu(E_{n+1}) + \sum_{i=1}^{n} \mu(E_i) \\
&= \sum_{i=1}^{n+1} \mu(E_i)
\end{aligned}$$

proving that $n+1\in\mathcal{S}$ $\square$

Now as $\mathcal{O}$ is a outer measure and $\bigsqcup_{i\in\mathbb{N}} E_i\in\mathcal{A}$ we have that $\mu(\bigsqcup_{i\in\mathbb{N}} E_i)=\mathcal{O}(\bigsqcup_{i\in\mathbb{N}} E_i)\leqslant\sum_{i=1}^{\infty}\mathcal{O}(E_i)=\sum_{i=1}^{\infty}\mu(E_i)$ proving that

$$\mu\left(\bigsqcup_{i\in\mathbb{N}} E_i\right)\leqslant\sum_{i=1}^{\infty}\mu(E_i) \tag{18.49}$$

Now given $n\in\mathbb{N}$ we have $\mu(\bigsqcup_{i\in\mathbb{N}} E_i)=\mathcal{O}(\bigsqcup_{i\in\mathbb{N}} E_i)\geqslant\mathcal{O}\big(\bigsqcup_{i\in\{1,\ldots,n\}} E_I\big)=\mu\big(\bigsqcup_{i\in\{1,\ldots,n\}} E_i\big)$ [as $\mathcal{O}$ is a outer measure and $\bigsqcup_{i\in\mathbb{N}} E_i\supseteq\bigsqcup_{i\in\{1,\ldots,n\}} E_i$). Hence using 18.48 we have that $\mu(\bigsqcup_{i\in\mathbb{N}} E_i)\geqslant\sum_{i\in\{1,\ldots,n\}}\mu(E_i)$. So $\mu(\bigsqcup_{i\in\mathbb{N}} E_i)\geq\sup\left(\{\sum_{i=1}^{n}\mu(E_i)|n\in\mathbb{N}\}\right)\underset{17.113}{=}\sum_{i=1}^{\infty}\mu(E_i)$. Applying then 18.49 we have finally have

$$\mu\left(\bigsqcup_{i\in\mathbb{N}}\mu(E_i)\right)=\sum_{i=1}^{\infty}\mu(E_i)$$

which together with 18.43 proves that μ is a measure. $\square$

As an example of the use of a outer measure we show how we can define a measure on a subset of a measure space. First we need a little lemma to create a outer measure based on a measure.

Lemma 18.35. *Let $\langle X,\mathcal{A},\mu\rangle$ be a measure space then for $\mu^*:\mathcal{P}(X)\to[0,\infty]$ defined by $\mu^*(A)=\inf(\{\mu(E)|E\in\mathcal{A}\wedge A\subseteq E\})$ we have [note that $0\leqslant\inf(\{\mu(E)|E\in\mathcal{A}\wedge A\subseteq E\})$ as 0 is a lower bound for $\{\mu(E)|E\in\mathcal{A}\wedge A\subseteq E\}$ and that the infinum exists as the real numbers are conditional complete (see 9.43)]*

1. *$\forall A\in\mathcal{P}(X)$ there exists a $E\in\mathcal{A}$ such that $\mu^*(A)=\mu(E)$ and $A\subseteq E$*
2. *μ^* is a outer measure*
3. *$\forall E\in\mathcal{A}$ we have $\mu^*(E)=\mu(E)$*

Proof. First as $\forall A\in\mathcal{P}(X)$ we have that $A\subseteq X\in\mathcal{P}(X)$ so that $\mu^*(A)=\{\mu(E)|E\in\mathcal{A}\wedge A\subseteq E\}\neq\emptyset$ and thus that $\inf(\{\mu(E)|E\in\mathcal{A}\wedge A\subseteq E\})$ exists.

1. Let $A\in\mathcal{P}(X)$ and $n\in\mathbb{N}$ then $\mu^*(A)<\mu^*(A)+\frac{1}{n}$ so using the definition of a infinum there exists a $E_n\in\mathcal{A}$ with $A\subseteq E_n$ such that

$$\mu^*(A)\leqslant\mu(E_n)<\mu^*(A)+\frac{1}{n} \tag{18.50}$$

Using 18.8 we have then that

$$E = \bigcap_{i \in \mathbb{N}} E_i \in \mathcal{A} \tag{18.51}$$

Now as $\forall n \in \mathbb{N}\ A \subseteq E_n$ it follows that $A \subseteq \bigcap_{i \in \mathbb{N}} E_i \in \mathcal{A}$ so that

$$\mu^*(A) \leqslant \mu(E) \tag{18.52}$$

Assume now that $\mu^*(A) < \mu(E)$ then $0 < \mu(E) - \mu^*(A)$ and using 9.55 there exists a $n \in \mathbb{N}$ such that $0 < \frac{1}{n} < \mu(E) - \mu^*(A)$ so that $\mu^*(A) < \mu^*(A) + \frac{1}{n} < \mu(E)$. applying 18.50 we have then that $\mu^*(A) < \mu(E_n) < \mu^*(A) + \frac{1}{n} < \mu(E)$ or $\mu(E_n) < \mu(E)$. As $E = \bigcap_{i \in \mathbb{N}} E_i \subseteq E_n$ we have $\mu(E) \leqslant \mu(E_n)$ leading to the contradiction $\mu(E_n) < \mu(E_n)$. Hence we must have

$$\mu^*(A) = \mu(E) \tag{18.53}$$

proving the first part of the lemma.

2. First note that $\emptyset \subseteq \emptyset \in \mathcal{A}$ so that $0 \leqslant \mu^*(\emptyset) \leqslant \{\mu(E) | E \in \mathcal{A} \wedge A \subseteq E\} \leqslant \mu(\emptyset) = 0$ proving that

$$\mu^*(\emptyset) = 0 \tag{18.54}$$

Take now $A, B \in \mathcal{P}(X)$ with $A \subseteq B$ if then if $x \in \{\mu(E) | E \in \mathcal{A} \wedge B \subseteq E\}$ there exists a $E \in \mathcal{A}$ such that $B \subseteq E$ and $x = \mu(E)$, as $A \subseteq B \subseteq E$ we have $x \in \{\mu(E) | E \in \mathcal{A} \wedge A \subseteq E\}$ proving $\{\mu(E) | E \in \mathcal{A} \wedge B \subseteq E\} \subseteq \{\mu(E) | E \in \mathcal{A} \wedge A \subseteq E\}$. Using 2.171 we have that $\inf(\{\mu(E) | E \in \mathcal{A} \wedge A \subseteq E\}) \leqslant \{\mu(E) | E \in \mathcal{A} \wedge B \subseteq E\}$ proving that

$$\forall A, B \in \mathcal{A} \text{ with } A \subseteq B \text{ we have } \mu^*(A) \leqslant \mu^*(B) \tag{18.55}$$

Let $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{P}(X)$ then $\forall i \in \mathbb{N}$ there exists a E_i such that $\mu^*(A_i) = \mu(E_i)$ and $A_i \subseteq E_i$ (see (1)). So $\bigcup_{i \in \mathbb{N}} A_i \subseteq \bigcup_{i \in \mathbb{N}} E_i \in \mathcal{A}$ [as $\mathcal{A}$ is a σ-algebra], hence $\mu^*(\bigcup_{i \in \mathbb{N}} A_i) = \inf(\{\mu(E) | E \in \mathcal{A} \wedge \bigcup_{i \in \mathbb{N}} A_i \subseteq E\}) \leqslant \mu(\bigcup_{i \in \mathbb{N}} E_i) \leqslant_{18.20} \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \mu^*(A_i)$ proving that

$$\forall \{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{P}(X) \text{ we have } \mu^*\left(\bigcup_{i \in \mathbb{N}} A_i\right) \leqslant \sum_{i=1}^{\infty} \mu(A_i) \tag{18.56}$$

Finally 18.54,18.55 and 18.56 proves part (2) of the theorem.

3. Let $E \in \mathcal{A}$ then using (1) there exists a $E' \in \mathcal{A}$ such that $E \subseteq E'$ and $\mu^*(E) = \mu(E')$. As $\mu^*(E) = \inf(\{\mu(F) | F \in \mathcal{A} \wedge E \subseteq F\})$ we have $\mu^*(E) \leqslant \mu(E)$, further from $E \subseteq E'$ we have $\mu(E) \leqslant \mu(E') = \mu^*(E)$. So $\mu^*(E) \leqslant \mu(E) \leqslant \mu^*(E)$ proving that $\mu(E) = \mu^*(E)$. $\square$

Theorem 18.36. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space and $A \subseteq X$ then $\mathcal{A}_A = \{A \bigcap E | E \in \mathcal{P}(X)\}$ is a σ-space and $\mu_A: \mathcal{A}_A \rightarrow [0, \infty]$ defined by $\mu_A(B) = \inf(\{\mu(E) | E \in \mathcal{A} \wedge B \subseteq E\})$*

1. *$\mathcal{A}_A$ is a σ-algebra on A*
2. *μ_A is a measure on $\mathcal{A}_A$*

*making $\langle A, \mathcal{A}_{|A}, \mu_A\rangle$ a measure space. μ_A is called the **subspace measure** on A. $\mathcal{A}_{|A}$ is called the **sub-algebra** induced by A (sometimes also called the **trace** algebra.*

Proof.

1. This follows from 18.9

2. We use the previous lemma (see 18.35) to get the outer-measure $\mu^*: \mathcal{P}(X) \rightarrow [0, \infty]$ defined by $\mu^*(B) = \inf(\{\mu(E) | E \in \mathcal{A} \wedge B \subseteq E\})$, using the definition of μ_A it follows then that $\mu_A = \mu^*_{|\mathcal{A}_A}$. First we have

$$\mu_A(\emptyset) = \mu^*(\emptyset) = 0 \tag{18.57}$$

Second let $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}_A$ be such that $\forall i, j \in \mathbb{N}$ we have $A_i \bigcap A_j = \emptyset$. First we have that as μ^* is a outer measure that

$$\mu^*\left(\bigcup_{i \in \mathbb{N}} A_i\right) \leqslant \sum_{i \in \mathbb{N}} \mu^\star(A_i) \tag{18.58}$$

For the opposite inequality note that by definition there exist $\forall i \in \mathbb{N}$ a $E_i \in \mathcal{A}$ such that $A_i = A \bigcap E_i$. Define now $B_i = \begin{cases} E_1 \text{ if } i = 1 \\ E_i \setminus \left(\bigcup_{j \in \{1, \ldots, i-1\}} A_j\right) \end{cases}$ then using 18.19 we have that $\{B_i\}_{i \in \mathbb{N}}$ are **pairwise disjoint** and that

$$\{B_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A} \text{ and } \bigsqcup_{i \in \mathbb{N}} B_i = \bigsqcup_{i \in \mathbb{N}} A_i \tag{18.59}$$

Further we have that

$$\forall i \in \mathbb{N} \text{ we have } A_i \subseteq B_i \tag{18.60}$$

Proof. Two cases must be considered

***i* = 1.** then $A_1 = A \bigcap E_1 \subseteq E_1 = B_1$ proving that $A_1 \subseteq B_1$

***i* > 1.** assume that $\exists j \in \{1, \ldots, i-1\}$ such that $x \in A_i \bigcap E_j$ then $x \in A \bigcap E_i \bigcap E_j = (A \bigcap E_i) \bigcap (A \bigcap E_j) = A_i \bigcap A_j \underset{i \neq j}{=} \emptyset$ a contradiction. So we have that $\forall j \in \{1, \ldots, i-1\}$ that $A_i \bigcap E_j = \emptyset$ hence $A_i \bigcap \left(\bigcup_{j \in \{1, \ldots, i-1\}} E_j\right) = \emptyset$. As $A_i = A \bigcap E_i \subseteq E_i$ we conclude that $A_i \subseteq E_i \setminus \left(\bigcup_{j \in \{1, \ldots, i-1\}} E_j\right)$ □

Now using 18.35 there exists a $E \in \mathcal{A}$ such that $\bigsqcup_{i \in \mathbb{N}} A_i \subseteq E$ and $\mu^*(\bigsqcup_{i \in \mathbb{N}} A_i) = \mu(E)$. Given $i \in \mathbb{N}$ define $C_i = B_i \bigcap E \in \mathcal{A}$ then $\{C_i\}_{i \in \mathbb{N}}$ is pairwise disjoint because $\{B_i\}_{i \in \mathbb{N}}$ is pairwise disjoint, also as $A_i \subseteq \bigcup_{i \in \mathbb{N}} A_i \subseteq E$ we have using 18.60 that $A_i \subseteq C_i$. So

$$\begin{aligned}
\sum_{i=1}^{\infty} \mu^*(A_i) &\leqslant \sum_{i=1}^{\infty} \mu^*(C_i) \\
&\underset{C_i \in \mathcal{A} \text{ and } 18.35\ (3)}{=} \sum_{i=1}^{\infty} \mu(C_i) \\
&\underset{\mu \text{ is a measure}}{=} \mu\left(\bigsqcup_{i \in \mathbb{N}} C_i\right) \\
&= \mu\left(\bigsqcup_{i \in \mathbb{N}} (B_i \bigcap E)\right) \\
&= \mu\left(E \bigcap \left(\bigsqcup_{i \in \mathbb{N}} B_i\right)\right) \\
&\underset{18.59}{=} \mu\left(E \bigcap \left(\bigsqcup_{i \in \mathbb{N}} A_i\right)\right) \\
&\underset{\bigsqcup_{i \in \mathbb{N}} A_i \subseteq E}{=} \mu(E) \\
&= \mu^*\left(\bigsqcup_{i \in \mathbb{N}} A_i\right)
\end{aligned}$$

which together with 18.58 proves that

$$\mu^*\left(\bigsqcup_{i \in \mathbb{N}} A_i\right) = \sum_{i=1}^{\infty} \mu^*(A_i) \tag{18.61}$$

Finally as $\mathcal{A}_A$ is a σ-algebra $\bigsqcup_{i \in \mathbb{N}} A_i \in \mathcal{A}_A$ so that

$$\begin{aligned}
\mu_A\left(\bigsqcup_{i \in \mathbb{N}} A_i\right) &= \mu^*\left(\bigsqcup_{i \in \mathbb{N}} A_i\right) \\
&\underset{18.61}{=} \sum_{i=1}^{\infty} \mu^*(A_i) \\
&\underset{A_i \in \mathcal{A}_A}{=} \sum_{i=1}^{\infty} \mu_A(A_i)
\end{aligned}$$

which together with 18.57 proves that μ_A is a measure. □

18.1.4.1 Lebesgue measure on $\mathbb{R}$

As an example of the Carathéodore construct we construct a measure on $\mathbb{R}$. We start with simple measurable objects, half open intervals where the length is a natural way to measure them.

Definition 18.37. *Given $a, b \in \mathbb{R}$ with $a \leqslant b$ then $[a, b[= \{x \in \mathbb{R} | a \leqslant x < b\}$ is called a half open interval and $[a, b] = \{x \in \mathbb{R} | a \leqslant x \leqslant b\}$ a closed interval. The set of all half open intervals is noted as $\mathcal{I}$ hence $\mathcal{I} = \{[a, b[| a, b \in \mathbb{R} \wedge a \leqslant b\}$*

Note 18.38. For $a \in \mathbb{R}$ we have $x \in [a, a[$ that $a \leqslant x < a$ giving the contradiction $a < a$ hence $[a, a[= \emptyset$ so $\emptyset = [1, 1[\in \mathcal{I}$.

To be able to define a length of a interval in a unique way we must first prove that the endpoints of a non empty half open intervals are unique.

Lemma 18.39. *Let $I \in \mathcal{I}$ with $I \neq \emptyset$ then there exists* ***unique*** *$a, b \in \mathbb{R}$ with $a < b$ such that $I = [a, b[$*

Proof. If $I \in \mathcal{I}$ then there exists a $a, b \in \mathbb{R}$ with $a \leqslant b$ such that $I = [a, b[$, as $I \neq \emptyset$ we must have that $a < b$. Assume that there exists a a', b' so that $[a, b[= [a', b'[$ then we have as $a \leqslant a < b$ that $a \in [a, b[= [a', b'[$ that

$$a \leqslant a' < b \tag{18.62}$$

If $a < a'$ then using 9.57 there exists a $x \in \mathbb{R}$ such that $a < x < a'$ and as $a' < b$ we have that $x \in [a, b[= [a', b'[$ hence $a' \leqslant x < b'$ giving the contradiction $a' < x < a'$. So we must have that $a' \leqslant a$ which using 18.62 proves

$$a = a' \tag{18.63}$$

If now $b < b'$ then using 9.57 again there exists a $x \in \mathbb{R}$ such that $b < x < b'$ and as $a' \underset{18.63}{=} a < b$ we have $a' < x < b'$ so that $x \in [a', b'[= [a, b[$ hence $x < b < b'$ contradicting $b < x < b'$. So we must have

$$b' \leqslant b \tag{18.64}$$

Assume now that $b' < b$ then using 9.57 there exists a $x \in \mathbb{R}$ such that $b' < x < b$ or as $a = a' \leqslant b'$ we have $x \in [a, b[= [a', b'[$ so that $x < b'$ contradicting $b' < x < b$. Hence $b \leqslant b'$ which using 18.64 prove that

$$b = b' \tag{18.65}$$

The lemma is then proved by 18.63 and 18.65 □

The above lemma ensures that the following definitions of begin,end and length of a half open interval is defined.

Definition 18.40. $\mathrm{begin}: \mathcal{I} \setminus \{\emptyset\} \to \mathbb{R}$ *and* $\mathrm{end}: \mathcal{I} \setminus \{\emptyset\} \to \mathbb{R}$ *are defined by* $\mathrm{begin}(I) = a$ *and* $\mathrm{end}(I) = b$ *where* $I = [a, b[$

Definition 18.41. *Let $I \in \mathcal{I}$ then $l(I) = \left\{ \begin{array}{l} 0 \text{ if } I \neq \emptyset \\ b - a \text{ where } I = [a, b[\end{array} \right.$*

Note 18.42. From the definition it easily follows that $\forall I \in \mathcal{I}$ we have that $0 \leqslant l(I)$

Lemma 18.43. *If $I \in \mathcal{I}$ and $\{I_i\}_{i \in \mathbb{N}} \subseteq \mathcal{I}$ is such that $I \subseteq \bigcup_{i \in \mathbb{N}} I_i$ then $\lambda(I) \leqslant \sum_{i=1}^{\infty} l(I_i)$*

Proof. We have the following cases to consider for I

$I = \emptyset$. then $l(I) = 0 \leqslant \sum_{i=1}^{\infty} l(I_i)$

I $\neq \emptyset$**.** then $\exists a,b\in\mathbb{R}$ with $a<b$ such that $I=[a,b[$, further for every $i\in\mathbb{N}$ there exists a $a_i,b_i\in\mathbb{R}$ such that $I_i=[a_i,b_i[$. Given $x\in\mathbb{R}$ define $\{I_i^x\}_{i\in\mathbb{N}}$ by $I_i^i=[a_i,\min(b_i,x)[$. Consider now the sets

$$A=\left\{x\in\mathbb{R}|x\in[a,b]\wedge x-a\leqslant\sum_{i=1}^{\infty}l(I_i^x)\right\}\subseteq[a,b] \tag{18.66}$$

Then as $a\in[a,b]$ and $a-a=0\leqslant\sum_{i=1}^{\infty}l(I_i^x)$ we have that $a\in A$ so that $\sup(A)$ exists. Further as $\forall x\in A$ we have $a\leqslant x\leqslant b$ we must have that $a\leqslant\sup(A)\leqslant b$. So if we define $c=\sup(A)$ we have

$$c=\sup(A)\text{ exists and } a\leqslant c\leqslant b \tag{18.67}$$

Define $B=\{\sum_{i=1}^{\infty}l(I_i^x)|x\in A\}$ then given $x\in A$ we have that $x-a\leqslant\sum_{i=1}^{\infty}l(I_i^x)\leqslant\sup(B)$ so that $x-a\leqslant\sup(B)$ or as a is finite that $x\leqslant\sup(B)+a$ hence $c=\sup(A)\leqslant\sup(B)+a$ or

$$c-a\leqslant\sup\left(\left\{\sum_{i=1}^{\infty}l(I_i^x)|x\in A\right\}\right) \tag{18.68}$$

Let $i\in\mathbb{N}$ then as $\forall x\in A$ we have $x\leqslant\sup(A)=c$ it follows $\min(b_i,x)\leqslant\min(b_i,c)$ so that $\min(b_i,x)-a_i\leqslant\min(b_i,c)-a_i$ or $l(I_i^x)\leqslant l(I_i^c)$. Hence

$$c-a\leqslant_{18.68}\sup\left(\left\{\sum_{i=1}^{\infty}l(I_i^x)|x\in A\right\}\right)\leqslant\sum_{i=1}^{\infty}l(I_i^c) \tag{18.69}$$

which as $c\in[a,b]$ (see 18.67) proves that

$$c\in A \tag{18.70}$$

Assume now that $c<b$ then $c\in[a,b[=\bigcup_{i\in\mathbb{N}}I_i$ so there exists a $k\in\mathbb{N}$ such that $c\in I_k=[a_k,b_k[\Rightarrow c<b_k$. Take $m=\min(b_k,b)$ then

$$c<m\text{ and } c<b_k \tag{18.71}$$

Now $\forall i\in\mathbb{N}$ we have $\min(b_i,c)\leqslant\min(b_i,m)$ or $\min(b_i,c)-a_i\leqslant\min(b_i,m)-a_i$ giving

$$\forall i\in\mathbb{N}\text{ we have } l(I_i^c)\leqslant l(I_i^m) \tag{18.72}$$

Further $l(I_k^m)=\min(b_k,m)-a_k\underset{m\leqslant b_k}{=}m-a_k$ and $l(I_k^c)=\min(c,b_k)-a_k\underset{c<b_k}{=}c-a_k$ hence

$$l(I_k^m)=l(I_k^c)+m-c \tag{18.73}$$

hence if we define $\{s_i\}_{i\in\mathbb{N}}$ by

$$s_i=\begin{cases} l(I_k^c)+m-c & \text{if } i=k\\ l(I_i^c)\text{ if } i\neq k\end{cases} \tag{18.74}$$

Using 18.72 and 18.73 we have that $\forall i\in\mathbb{N}$ that $l(I_i^m)\geqslant c_i$ so using 17.115 we have that

$$\sum_{i=1}^{\infty}l(I_i^m)\geqslant\sum_{i=1}^{\infty}s_i \tag{18.75}$$

Now if we define $\{a_i\}_{i\in\mathbb{N}}$ by $a_i = \left\{ \begin{smallmatrix} m - c \text{ if } i = k \\ 0 \text{ if } i \neq k \end{smallmatrix} \right.$ then using 18.74 we have that $\forall i \in \mathbb{N}\, s_i = l(I_i^c) + a_i$ so that $\sum_{i=1}^{\infty} s_i = \sum_{i=1}^{\infty} (l(I_i^c) + a_i) \underset{17.114}{=} \sum_{i=1}^{\infty} l(I_i^c) + \sum_{i=1}^{\infty} a_i = \sum_{i=1}^{\infty} l(I_i^c) + m - c$. Hence using 18.75 we have $\sum_{i=1}^{\infty} l(I_i^m) \geqslant \sum_{i=1}^{\infty} l(I_i^c) + m - c$. Applying then 18.69 we have

$$\sum_{i=1}^{\infty} l(I_i^m) \geqslant c - a + m - c = m - a \tag{18.76}$$

As we also have that $c \in [a_k, b_k[\Rightarrow c < b_k$ and $a \leqslant c$ it follow that $a \leqslant \min(b, b_k) \leqslant b \Rightarrow m = \min(b, b_k) \in [a, b]$ which together with 18.76 proves that $m \in A$ hence $m \leqslant \sup(A) = c$ contradicting 18.71. So the assumption $c < b$ is false and we must have that $b \leqslant c$, using 18.67 we have $c = b$ or $b \in A$, which by the definition of A gives $b - a \leqslant \sum_{i=1}^{\infty} l(I_i^b)$. As $\forall i \in \mathbb{N}$ we have that $l(I_i^b) = \min(b, b_i) - a_i \leqslant b_i - a_i = l(I_i)$ it follows that

$$l(I) = b - a \leqslant \sum_{i=1}^{\infty} l(I_i) \tag{18.77}$$

proving the theorem. □

To be able to apply the Carathédory theorem we must now define a outer measure. To do this we must define a outer measure on $\mathbb{R}$. This is done in the following lemma:

Lemma 18.44. *$\varphi: \mathcal{P}(X) \rightarrow [0, \infty]$ where $\varphi(A) = \inf(\{\sum_{i=1}^{\infty} l(I_i) | \{I_i\}_{i\in\mathbb{N}} \subseteq \mathcal{I} \text{ such that } A \subseteq \bigcup_{i\in\mathbb{N}} I_i\})$ is a outer measure on $\mathbb{R}$. Further $\forall I \in \mathcal{I}$ we have $\varphi(I) = l(I)$.*

Proof. We have

1. $\emptyset \in \mathcal{I}$ as $[1, 1[$
2. If $x \in \mathbb{R}$ then using a consequence of the Archimedean property of the real numbers (see 9.55) there exists a $n \in \mathbb{N}$ such that $|x| < n$ then $x < n$ and $-x < n \Rightarrow -n < x$ proving that $x \in [-n, n[\subseteq \bigcup_{n\in\mathbb{N}} [-n, n[$. So $\mathbb{R} \subseteq \bigcup_{n\in\mathbb{N}} [-n, n[$
3. $l(\emptyset) = 0$
4. $\forall I \in \mathcal{I}, \{I_i\}_{i\in\mathbb{N}} \subseteq \mathcal{I}$ with $I \subseteq \bigcup_{i\in\mathbb{N}} I_i$ we have $l(I) \leqslant \sum_{i=1}^{\infty} l(I_i)$ (see the previous lemma 18.43)

So all the conditions of theorem 18.33 are satisfied which proves the theorem. □

We are now ready to define a σ-algebra and a measure on $\mathbb{R}$ using the Carathéodory theorem (see 18.34)

Definition 18.45. *Let $\varphi(A) = \inf(\{\sum_{i=1}^{\infty} l(I_i) | \{I_i\}_{i\in\mathbb{N}} \subseteq \mathcal{I} \text{ where } A \subseteq \bigcup_{i\in\mathbb{N}} I_i\})$ be the outer measure from the previous lemma then the σ-algebra defined by the Carathéodory method $\{E \in \mathcal{P}(X) | \forall A \in \mathcal{P}(X) \text{ we have } \varphi(A) = \varphi(A \bigcap E) + \varphi(A \setminus E)\}$ is called the set of Lebesgue measurable sets and is noted by $\mathcal{L}$. The measure λ: $\mathcal{L} \rightarrow [0, \infty]$ defined by $\lambda = \varphi_{|\mathcal{L}}$ is called the Lebesgue measure. This makes $\langle \mathbb{R}, \mathcal{L}, \lambda \rangle$ a measure space.*

We prove now that the Borel algebra $\mathcal{B}[\mathbb{R}]$ (see 18.12) is a subset of $\mathcal{L}$ or in other words that all Borel measurable sets on $\mathbb{R}$ are Lebesgue measurable. First we prove that segments of the form $]-\infty, x[$ are Lebesgue measurable. Before that we have a little lemma about the intersection of a segment and a interval.

Lemma 18.46. *Let $x \in \mathbb{R}$ and $I = [a, b[\in \mathcal{I}$ then we have for $I \bigcap]-\infty, x[$ and $I \setminus]-\infty, x[$ either*

1. $I \bigcap]-\infty, x[= \emptyset \wedge I \setminus]-\infty, x[= I$
2. $I \bigcap]-\infty, x[= I \wedge I \setminus]-\infty, x[= \emptyset$
3. $I \bigcap]-\infty, x[= [a, x[\wedge I \setminus]-\infty, x[= [x, b[$

In other words $I \bigcap]-\infty, x[\in \mathcal{I}$ and $I \setminus]-\infty, x[\in \mathcal{I}$

Proof. As $I \in \mathcal{I}$ there exists a $a, b \in \mathbb{R}$ such that $a \leqslant b$ and $I = [a, b[$. Then for a, b, x we have the following cases to consider

$\boldsymbol{a = b}$**.** then of course $I = \emptyset$ so that $I \bigcap]-\infty, x[= \emptyset$ and $I \setminus]-\infty, x[= \emptyset = I$ giving (1) of the lemma.

$\boldsymbol{a < b}$**.**

$\boldsymbol{a < b \leqslant x}$**.** then if $y \in I$ we have $x \leqslant y < b \leqslant x$ so that $y \in]-\infty, x[$ proving that $I \subseteq]-\infty, x[$. Hence $I \bigcap]-\infty, x[= I$ and $I \setminus]-\infty, x[= \emptyset$ giving (2) of the lemma.

$\boldsymbol{a \leqslant x < b}$**.** then

$$\begin{aligned} y \in I \bigcap]-\infty, x[&\Leftrightarrow a \leqslant y < b \wedge y < x \\ &\underset{\mathrm{x<b}}{\Leftrightarrow} a \leqslant y < x \\ &\Leftrightarrow y \in [a, x[\end{aligned} \tag{18.78}$$

$$\begin{aligned} y \in I \setminus]-\infty, x[&\Leftrightarrow a \leqslant y < b \wedge \neg(y < x) \\ &\Leftrightarrow a \leqslant y < b \wedge x \leqslant y \\ &\underset{a \leqslant x}{\Leftrightarrow} x \leqslant y < b \\ &\Leftrightarrow y \in [x, b[\end{aligned} \tag{18.79}$$

giving (3) of the lemma.

$\boldsymbol{x < a < b}$**.** then if $y \in I \bigcap]-\infty, x[$ then $a \leqslant y < b$ and $y < x < a$ giving the contradiction $y < y$ proving that $I \bigcap]-\infty, x[= \emptyset$. From this it follows also that $I \setminus]-\infty, x[= I$ giving (1) of the lemma. □

Lemma 18.47. *Let $x \in \mathbb{R}$ then $]-\infty, x[$ is Lebesgue measurable*

Proof. First we prove that for the φ in the above definition

$$\forall I \in \mathcal{I} \text{ we have } \varphi(I) = \varphi\left(I \bigcap]-\infty, x[\right) + \varphi(I \setminus]-\infty, x[) \tag{18.80}$$

Proof. As $I \in \mathcal{I}$ then there exists a $a, b \in \mathbb{R}$ with $a \leqslant b$ and $I = [a, b[$, we have then using 18.46 the following cases to consider for $I \bigcap]-\infty, x[$ and $I \setminus]-\infty, x[$

$\boldsymbol{I \bigcap]-\infty, x[= \emptyset \wedge I \setminus]-\infty, x[= I}$**.** then $\varphi(I \bigcap]-\infty, x[) + \varphi(I \setminus]-\infty, x[) = \varphi(\emptyset) + \varphi(I) = 0 + \varphi(I) = \varphi(I)$

$\boldsymbol{I\bigcap\,]-\infty,x[\,=I\wedge I\setminus]-\infty,x[\,=\emptyset}$**.** then $\varphi(I\bigcap\,]-\infty,x[)+\varphi(I\setminus]-\infty,x[)=\varphi(I)+\varphi(\emptyset)=\varphi(I)$

$\boldsymbol{I\bigcap\,]-\infty,x[\,=[a,x[\,\wedge I\setminus]-\infty,x[\,=[x,b[}$**.** then

$$\begin{aligned}\varphi\left(I\bigcap\,]-\infty,x[\right)+\varphi(I\setminus]-\infty,x[) &= \varphi([a,x[)+\varphi([x,b[)\\ &\underset{18.44}{=} \lambda([a,x[)+\lambda([x,b[)\\ &= x-a+b-x\\ &= b-a\\ &= \lambda(I)\\ &\underset{18.44}{=} \varphi(I)\end{aligned}$$

So in all cases we have $\varphi(I\bigcap\,]-\infty,x[)+\varphi(I\setminus]-\infty,x[)=\varphi(I)$ □

Take now $A\in\mathcal{P}(X)$ then by the definition of φ (see 18.44) we have $\forall\varepsilon>0$ that there exists a $\{I_i\}_{i\in\mathbb{N}}\subseteq\mathcal{I}$ such that $A\subseteq\bigcup_{i\in\mathbb{N}}A_i$ and

$$\varphi(A)\leqslant\sum_{i=1}^{\infty}l(I_i)<\varphi(A)+\varepsilon \tag{18.81}$$

Now $\forall i\in\mathbb{N}$ we have that $I_i\bigcap\,]-\infty,x[\,\in\mathcal{I}\wedge I_i\setminus]-\infty,x[\,\in\mathcal{I}$ (see 18.46). Now as $A\bigcap\,]-\infty,x[\,\subseteq(\bigcup_{i\in\mathbb{N}}I_i)\bigcap\,]-\infty,x[\,\underset{1.107}{=}\bigcup_{i\in\mathbb{N}}(I_i\bigcap\,]-\infty,x[)$ and $A\setminus]-\infty,x[\,=(\bigcup_{i\in\mathbb{N}}I_i)\setminus\,]-\infty,x[\,\underset{1.108}{=}\bigcup_{i\in\mathbb{N}}(I_i\setminus]-\infty,x[)$ hence we have using the definition of φ that $\varphi(A\bigcap\,]-\infty,x[)\leqslant\sum_{i=1}^{\infty}\varphi(I_i\bigcap\,]-\infty,x[)$ and $\varphi(A\setminus]-\infty,x[)\leqslant\sum_{i=1}^{\infty}\varphi(I_i\setminus]-\infty,x[)$ so that

$$\begin{aligned}\varphi\left(A\bigcap\,]-\infty,x[\right)+\varphi(A\setminus]-\infty,x[) &\leqslant \sum_{i=1}^{\infty}\varphi\left(I_i\bigcap\,]-\infty,x[\right)+\sum_{i=1}^{\infty}\varphi(I_i\setminus]-\infty,x[)\\ &\underset{17.114}{=} \sum_{i=1}^{\infty}\left(\varphi\left(I_i\bigcap\,]-\infty,x]\right)+\varphi(I_i\setminus]-\infty,x[)\right)\\ &\underset{18.79}{=} \sum_{i=1}^{\infty}\varphi(I_i)\\ &\leqslant_{18.81} \varphi(A)+\varepsilon\end{aligned}$$

As ε is chosen arbitrary we can use 9.56 to prove that $\forall A\in\mathcal{P}(X)$ we have $\varphi(A\bigcap\,]-\infty,x[)+\varphi(A\setminus]-\infty,x[)\leqslant\varphi(A)$. Hence $]-\infty,x[\,\in\{E\in\mathcal{P}(X)|\forall A\in\mathcal{P}(X)$ we have $\varphi(A)\geqslant\varphi(A\bigcap E)+\varphi(A\setminus E)\}\underset{18.34}{=}\{E\in\mathcal{P}(X)|\forall A\in\mathcal{P}(X)$ we have $\varphi(A)=\varphi(A\bigcap E)+\varphi(A\setminus E)\}=\mathcal{L}$ proving that $]-\infty,x[$ is Lebesgue measurable. □

A consequence of the above is that the σ-algebra generated by a segments of the form $]-\infty,x[$ is a subset of the set of Lebesgue measurable

Corollary 18.48. $\sigma[\{]-\infty,x[|x\in\mathbb{R}\}]\subseteq\mathcal{L}$

Proof. Using the previous lemma 18.47 we have that $\{]-\infty,x[|x\in\mathbb{R}\}\subseteq\mathcal{L}$. As $\mathcal{L}$ is a σ-algebra and $\sigma[\{]-\infty,x[|x\in\mathbb{R}\}]$ is the smallest σ-algebra covering $\{]-\infty,x[|x\in\mathbb{R}\}\subseteq\mathcal{L}$. As $\mathcal{L}$ is a σ-algebra and $\sigma[\{]-\infty,x[|x\in\mathbb{R}\}]$ is the smallest σ-algebra covering $\{]-\infty,x[|x\in\mathbb{R}\}$ (see 18.10) it follows that $\sigma(\{]-\infty,x[|x\in\mathbb{R}\})$ □

Corollary 18.49. $\sigma[\mathcal{I}] = \sigma[\{]-\infty, x[|x \in \mathbb{R}\}] \subseteq_{18.48} \mathcal{L}$

Proof. Let $I \in \mathcal{I}$ then there exists a $a, b \in \mathbb{R}$ such that $I = [a, b[$. Now

$$\begin{aligned} x \in]-\infty, b[\setminus]-\infty, a[&\Leftrightarrow x < b \wedge \neg(x < a) \\ &\Leftrightarrow x < b \wedge a \leqslant x \\ &\Leftrightarrow x \in [a, b[\end{aligned} \tag{18.82}$$

As $\sigma[\{]-\infty, x[|x \in \mathbb{R}\}]$ is a σ-algebra we have using 18.8 that $]-\infty, b[\setminus]-\infty, a[\in \sigma[\{]-\infty, x[|x \in \mathbb{R}\}]$ so that by 18.82 it follows that $I = [a, b[\in \sigma[\{]-\infty, x[|x \in \mathbb{R}\}]$. Hence $\mathcal{I} \subseteq \sigma[\{]-\infty, x[|x \in \mathbb{R}\}]$ so that $\sigma[\mathcal{I}] \subseteq \sigma[\{]-\infty, x[|x \in \mathbb{R}\}]$ (see 18.10) ☐

We can extend the above to open sets. to do this we first show that every open set can be written as a countable union of half open intervals. Actually we go even further and show that it can be written as a countable disjoint union of intervals. For this we introduce the concept of Dyadic intervals.

Definition 18.50. *Let $n \in \mathbb{N}_0$, $z \in \mathbb{Z}$ then $\left[\frac{z}{2^n}, \frac{z+1}{2^n}\right[$ is called a **Dyadic** interval of order n. The set of all Dyadic intervals of order n is noted by $\mathcal{D}_n$ so $\mathcal{D}_n = \left\{\left[\frac{z}{2^n}, \frac{z+1}{2^n}\right[|z \in \mathbb{Z}\right\}$. The set of all Dyadic intervals is noted by $\mathcal{D}$ so $\mathcal{D} = \bigcup_{n \in \mathbb{N}_0} \mathcal{D}_n$.*

Note 18.51. As $\frac{z}{2^n} < \frac{z+1}{2^n}$ we have that Dyadic intervals are non empty.

We prove now that the set of Dyadic interval is denumerable

Lemma 18.52. $\forall n \in \mathbb{N}$ *$\mathcal{D}_n$ is denumerable and in addition $\mathcal{D}$ is also denumerable.*

Proof. Let $n \in \mathbb{N}_0$ define then $\beta: \mathbb{Z} \rightarrow \mathbb{D}_n$ by $\beta(z) = \left[\frac{z}{2^n}, \frac{z+1}{2^n}\right[$ then β is a bijection:

injectivity. If $\beta(x) = \beta(y)$ then $\left[\frac{x}{2^n}, \frac{x+1}{2^n}\right[= \left[\frac{y}{2^n}, \frac{y+1}{2^n}\right[\underset{18.39}{\Rightarrow} \frac{x}{2^n} = \frac{y}{2^n} \Rightarrow x = y$

surjectivity. If $I \in \mathbb{D}_n$ then $\exists z \in \mathbb{Z}$ so that $I = \left[\frac{z}{2^n}, \frac{z+1}{2^n}\right[= \beta(z)$

As $\mathbb{Z}$ is denumerable (see 8.54) we have thus that $\mathcal{D}_n$ is denumerable. Using 5.60 we have then that $\mathcal{D} = \bigcup_{i \in \mathbb{N}_0} \mathcal{D}_i$ is also denumerable. ☐

Dyadic intervals have the following important property

Lemma 18.53. *Let $n, k \in \mathbb{N}_0$ with $k \leqslant n$ and $v, z \in \mathbb{Z}$ then we have either $\left[\frac{v}{2^n}, \frac{v+1}{2^n}\right[\bigcap \left[\frac{z}{2^k}, \frac{z+1}{2^k}\right[= \emptyset$ or $\left[\frac{v}{2^n}, \frac{v+1}{2^n}\right[\subseteq \left[\frac{z}{2^k}, \frac{z+1}{2^k}\right[$*

Proof. Take $I = \left[\frac{v}{2^n}, \frac{v+1}{2^n}\right[$ and $J = \left[\frac{z}{2^k}, \frac{z+1}{2^k}\right[$. As $k < n$ we have $0 \leqslant l = n - k$ and

$$\begin{aligned} J &= \left[\frac{z}{2^k}, \frac{z+1}{2^k}\right[\\ &= \left[\frac{z}{2^k} \cdot \frac{2^{n-k}}{2^{n-k}}, \frac{z \cdot 2^{n-k} + 2^{n-k}}{2^k \cdot 2^{n-k}}\right[\\ &= \left[\frac{z \cdot 2^{n-k}}{2^n}, \frac{z \cdot 2^{n-k} + 2^{n-k}}{2^n}\right[\\ &= \left[\frac{z \cdot 2^l}{2^n}, \frac{z \cdot 2^l + 2^l}{2^n}\right[\end{aligned}$$

proving that

$$J = \left[\frac{z \cdot 2^l}{2^n}, \frac{z \cdot 2^l + 2^l}{2^n}\right[\tag{18.83}$$

If now $I \bigcap J \neq 0$ so there exists a $x \in I \bigcap J$ then we have if

$\boldsymbol{\frac{v+1}{2^n} \leqslant \frac{z \cdot 2^l}{2^n}}$. then $\frac{v}{2^n} \leqslant x < \frac{v+1}{2^n} \wedge \frac{v+1}{2^n} \leqslant \frac{z \cdot 2^l}{2^n} \leqslant x < \frac{z \cdot 2^l + 2^l}{2^n}$ giving $x < x$ a contradiction.

$\boldsymbol{\frac{z \cdot 2^l + 2^l}{2^n} \leqslant \frac{v}{2^n}}$. then $\frac{z \cdot 2^l}{2^n} \leqslant x < \frac{z \cdot 2^l + 2^l}{2^n} \wedge \frac{z \cdot 2^l + 2^l}{2^n} \leqslant \frac{v}{2^n} \leqslant x < \frac{v+1}{2^n}$ giving $x < x$ a contradiction

so we must have $\frac{z \cdot 2^l}{2^n} < \frac{v+1}{2^n}$ and $\frac{v}{2^n} < \frac{z \cdot 2^l + 2^l}{2^n}$ so by multiplication with 2^n gives

$$z \cdot 2^l < v + 1 \wedge v < z \cdot 2^l + 2^l \tag{18.84}$$

Assume now that $\frac{v}{2^n} < \frac{z \cdot 2^l}{2^n}$ then $v < z \cdot 2^l$ and as $v, z \cdot 2^l \in \mathbb{Z}$ we have $v + 1 \leqslant z \cdot 2^l <_{18.84} v + 1$ a contradiction so we must have

$$\frac{z \cdot 2^l}{2^n} \leqslant \frac{v}{2^n} \tag{18.85}$$

Assume that $\frac{z \cdot 2^l + 2^l}{2^n} < \frac{v+1}{2^n}$ then $z \cdot 2^l + 2^l < v + 1$ so as $\cdot 2^l + 2^l, v + 1 \in \mathbb{Z}$ we have $z \cdot 2^l + 2^l \leqslant v$ contradicting 18.84 so we must have

$$\frac{v+1}{2^n} \leqslant \frac{z \cdot 2^l + 2^l}{2^n} \tag{18.86}$$

So if $x \in I$ then $\frac{v}{2^n} \leqslant x < \frac{v+1}{2^n}$ and using 18.85 and 18.86 we have $\frac{z \cdot 2^l}{2^n} \leqslant x < \frac{z \cdot 2^l + 2^l}{2^n}$ giving by 18.83 that $x \in I$. So we have that if $I \bigcap J \neq \emptyset$ then $I \subseteq J$. Further if $I \bigcap J = \emptyset$ then if $I \subseteq J$ we have as $I \neq \emptyset$ that $\exists x \in I \subseteq J$ giving $I \bigcap J \neq \emptyset$ a contradiction. Hence we must either have that $I \bigcap J = \emptyset$ or $I \subseteq J$ but not both $I \bigcap J = \emptyset$ and $I \subseteq J$. □

We prove now the extra properties of Dyadic sets

Lemma 18.54. *Dyadic intervals have the following properties*

1. *If $k, n \in \mathbb{N}_0$ with $k \leqslant n$ we have $\forall I \in \mathcal{D}_n, \forall J \in \mathcal{D}_k$ then either $I \subseteq J$ or $I \bigcap J = \emptyset$*
2. *$\forall n \in \mathbb{N}_0$ then $\forall I, J \in \mathcal{D}_n$ with $I \neq J$ we have $I \bigcap J = \emptyset$*
3. *$\forall n \in \mathbb{N}_0$ we have $\mathbb{R} = \bigsqcup_{I \in \mathcal{D}_n} I$*
4. *$\forall n \in \mathbb{N}_0$ and $\forall I \in \mathcal{D}_n$ we have $l(I) = \frac{1}{2^n}$*

Proof.

1. This follows directly from 18.53

2. Take $I \neq J$ and assume that $I \bigcap J \neq \emptyset$ then as $n \leqslant n$ we have by (1) that $I \subseteq J \wedge J \subseteq I \Rightarrow I = J$ a contradiction, so we must have that $I \bigcap J = \emptyset$

3. Let $n \in \mathbb{N}_0$ then trivially $\bigcup_{I \in \mathbb{D}_n} I \subseteq \mathbb{R}$, for the opposite inclusion let $x \in \mathbb{R}$ then for $x \cdot 2^n$ we have by the Archimedean property of the reals (see 9.55) that there exists a $z \in \mathbb{Z}$ such that $z \leqslant x \cdot 2^n < z + 1$ and thus that $\frac{z}{2^n} \leqslant x < \frac{z+1}{2^n} \Rightarrow x \in \left[\frac{z}{2^n}, \frac{z+1}{2^n}\right[$ proving that $\mathbb{R} \subseteq \bigcup_{I \in \mathbb{D}_n} I$. Using (2) proves then that $\mathbb{R} = \bigsqcup_{I \in \mathbb{D}_n} I$

4. If $I \in \mathbb{D}_n$ then $\exists z \in \mathbb{Z}$ such that $I = \left[\frac{z}{2^n}, \frac{z+1}{2^n}\right[$ so that $l(I) = \frac{z+1}{2^n} - \frac{z}{2^n} = \frac{1}{2^n} = 2^{-n}$ □

We can now use Dyadic intervals to show that every open set is a disjoint union of Dyadic intervals.

Lemma 18.55. *Let $\mathbb{R}$ equipped with the canonical topology $\mathcal{T}_{\mathbb{R}}$ then we have $\forall U \in \mathcal{T}_{\mathbb{R}}$ with $\emptyset \neq U$ that there exists a sequence of pairwise disjoint Dyadic intervals $\{I_i\}_{i \in \mathbb{N}} \subseteq \mathcal{D}$ such that $U = \bigsqcup_{i \in \mathbb{N}} I_i$*

Proof. Let U be a non empty open set in $\mathbb{R}$, define then

$$\{S_n\}_{n \in \mathbb{N}} \text{ by } S_n = \{I \in \mathcal{D}_n | I \subseteq U\} \subseteq \mathcal{D}_n \tag{18.87}$$

Further define recursively

$$\{\mathcal{T}_n\}_{n \in \mathbb{N}_0} \text{ by } \mathcal{T}_n = \left\{ \begin{array}{l} \mathcal{S}_1 \text{ if } n = 1 \\ \{I \in \mathcal{S}_n | \forall i \in \{1, \ldots, n-1\} \text{ we have } \forall I \in \mathcal{T}_i \vDash I \bigcap J = \emptyset\} \end{array} \right. \subseteq \mathcal{S}_n \subseteq \mathcal{D}_n \tag{18.88}$$

Take now

$$\mathcal{T} = \bigcup_{i \in \mathbb{N}} \mathcal{T}_i \subseteq \bigcup_{i \in \mathbb{N}} \mathcal{D}_n \subseteq \mathcal{D} \tag{18.89}$$

Let $x \in U$ then as U is open there exists a $0 < \varepsilon_x$ such that $x \in]x - \varepsilon_x, x + \varepsilon_x[\subseteq U$. Using a consequence of the Archimedean property of the real numbers (see 9.66) there exists a $n_x \in \mathbb{N}$ such that

$$0 < \frac{1}{2^{n_x}} < \varepsilon \tag{18.90}$$

Using 18.54 (3) there exists a $I_x \in \mathcal{D}_{n_x}$ such that $x \in I_x$, hence there exists a $z \in \mathbb{Z}$ such that $I_x = \left[\frac{z}{2^{n_x}}, \frac{z+1}{2^{n_x}}\right[$ proving

$$\frac{z}{2^{n_x}} \leqslant x < \frac{z+1}{2^{n_x}} \tag{18.91}$$

and

$$\forall t \in I_x \text{ we have } \frac{z}{2^{n_x}} \leqslant t < \frac{z+1}{2^{n_x}} \tag{18.92}$$

Hence

$$\begin{array}{rcl} x-t & \leqslant_{18.91,18.92} & \frac{z+1}{2^{n_x}}-\frac{z}{2^{n_x}} \\ & = & \frac{1}{2^{n_x}} \\ & < & \varepsilon_x \\ t-x & \leqslant_{18.91,18.92} & \frac{z+1}{2^{n_x}}-\frac{z}{2^{n_x}} \\ & = & \frac{1}{2^{n_x}} \\ & < & \varepsilon_x \end{array}$$

hence $x-\varepsilon_x<t<x+\varepsilon_x$ proving that $t\in]x-\varepsilon_x,x+\varepsilon_x[\subseteq U$. So we have that $I_x\subseteq U$ or using 18.87 that $x\in I_x\in\mathcal{S}_{n_x}$. Define now $\mathcal{K}_x=\{n\in\{1,\ldots,n_x\}|\exists I\in\mathcal{S}_n \text{such that} x\in I\}$ then $n_x\in\mathcal{K}_x$ proving that $\mathcal{K}_x\neq\emptyset$ so that $m_x=\min(\mathcal{K}_x)$ exists. We have then two cases to consider

$m_x=1$. then $\exists I\in S_1=\mathcal{T}_1\in\mathcal{T}$ such that $x\in I$ hence $x\in I\subseteq\bigcup_{I\in\mathcal{T}}I$

$1<m_x$. then $\exists J\in\mathcal{S}_{m_x}$ (such that $x\in J$) and for $n\in\{1,\ldots,m_x-1\}$ we have $\forall I\in\mathcal{S}_n$ that $x\notin I$. Now as $n\leqslant m_x$ we have by 18.54 (1) that either $J\subseteq I$, but then as $x\in J\subseteq I$ we contradict$x\notin I$, or $J\bigcap I=\emptyset$. So $\forall n\in\{1,\ldots,m_x-1\}$, $\forall I\in\mathcal{I}_n$ we have $J\bigcap I=\emptyset$ or $J\in\mathcal{T}_{m_x}$ (see 18.88) hence $x\in J\subseteq\bigcup_{I\in\mathcal{T}}I$

As we have chosen $x\in U$ arbitrary it follows that

$$U\subseteq\bigcup_{I\in\mathcal{T}}I \tag{18.93}$$

Now if $x\in\bigcup_{I\in\mathcal{T}}I$ then there exists a $I\in\mathcal{T}$ such that $x\in I$, as $\mathcal{T}=\bigcup_{i\in\mathbb{N}_0}\mathcal{T}_i$ there exists a $i\in\mathbb{N}$ such that $I\in\mathcal{T}_i$ and thus by 18.88 that $I\subseteq U$. So $\bigcup_{I\in\mathcal{T}}I\subseteq U$ which by 18.93 proves that

$$U=\bigcup_{I\in\mathcal{T}}I \tag{18.94}$$

Now if $I,J\in\mathcal{T}$ with $I\neq J$ then there exists n,m such that $I\in\mathcal{T}_n$ and $J\in\mathcal{T}_m$. For n,m we have the following possibilities:

n=m. then using 18.54 (2) we have that $I\bigcap J=\emptyset$

$n<m$. then using 18.88 we have $I\bigcap J=\emptyset$

$m<n$. then using 18.88 we have $I\bigcap J=\emptyset$

Applying this on 18.94 proves that

$$U=\bigsqcup_{I\in\mathcal{T}}I \tag{18.95}$$

Assume now that $\mathcal{T}$ is finite, then as $U \neq \emptyset$ we must have that $\mathcal{T} \neq \emptyset$. As Dyadic intervals are non empty we have that $a = \min(\{\text{begin}(I) | I \in \mathcal{T}\})$ exists. By the definition of a minimum there exists a $I \in \mathcal{T}$ such that $\text{begin}(I) = a$ hence $a \in I \subseteq U$. As U is open there exists a $\varepsilon > 0$ such that $a \in]a - \varepsilon, a + \varepsilon[\subseteq U$ hence $a - \frac{\varepsilon}{2} \in U$. So there exists a $J \in \mathcal{T}$ such that $a - \frac{\varepsilon}{2} \in J$ hence $\text{begin}(J) \leqslant a - \frac{\varepsilon}{2} < a \leqslant \text{begin}(J)$ a contradiction. So we must have that $\mathcal{T}$ is infinite. As $\mathcal{T} \subseteq \mathcal{D}$ a denumerable set (see 18.52) it follows from 5.55 that

$$\mathcal{T} \text{ is denumerable} \tag{18.96}$$

So there exists a bijection $\beta: \mathbb{N} \to \mathcal{T}$, define then $\{I_i\}_{I \in \mathbb{N}}$ by $I_i = \beta(i)$ giving

$$U = \bigsqcup_{i \in \mathbb{N}} I_i$$

proving the lemma. □

Lemma 18.56. *$\forall I \in \mathcal{I}$ with $\emptyset \neq I$ there exists a sequence $\{U_i\}_{i \in \mathbb{N}} \subseteq \mathcal{T}_{\mathbb{R}}$ of open sets such that $\bigcap_{i \in \mathbb{N}} U_i = I$*

Proof. As $I \neq \emptyset$ there exists a $a, b \in \mathbb{R}$ with $a < b$ such that $I = [a, b[$ define then $\{]a - \frac{1}{n}, b[\}_{n \in \mathbb{N}} \subseteq \mathcal{T}_{\mathbb{R}}$. As $\forall n \in \mathbb{N}$ $[a, b[\subseteq]a + \frac{1}{n}, b[$ we have

$$I \subseteq \bigcap_{n \in \mathbb{N}} \left] a - \frac{1}{n}, b \right[\tag{18.97}$$

For the opposite inequality let $x \in \bigcap_{n \in \mathbb{N}}]a - \frac{1}{n}, b[$ then $\forall n \in \mathbb{N}$ we have $a - \frac{1}{n} \leqslant x < b$. Assume now that $x < a$ then $0 < a - x$, so that using 9.55 there exists a $n \in \mathbb{N}$ such that $0 < \frac{1}{n} < a - x \Rightarrow x < a - \frac{1}{n}$ contradicting $a - \frac{1}{n} < x$, so $a \leqslant x < b$ or $x \in [a, b[$. Hence $\bigcap_{n \in \mathbb{N}}]a - \frac{1}{n}, b] \subseteq I$ proving using 18.97 that

$$I = \bigcap_{i \in \mathbb{N}} \left] a - \frac{1}{n}, b \right[$$

□

We are now ready to prove that $\mathcal{B}[\mathbb{R}] = \sigma[\mathcal{I}]$

Lemma 18.57. *$\mathcal{B}[\mathbb{R}] = \sigma[\mathcal{I}]$ (where $\mathcal{B}[\mathbb{R}] = \mathcal{B}[\mathbb{R}, \mathcal{T}_{\mathbb{R}}] = \sigma[\mathcal{T}_{\mathbb{R}}]$ see 18.13)*

Proof. Let $U \in \mathcal{T}_{\mathbb{R}}$ then we have either

- ***$U = \emptyset$.*** then $U \in \sigma[\mathcal{I}]$
- ***$U \neq \emptyset$.*** then using 18.55 there exists a $\{I_i\}_{i \in \mathbb{N}} \subseteq \mathcal{I}$ such that $U = \bigcup_{i \in \mathbb{N}} I_i \in \sigma[\mathcal{I}]$ [as $\sigma[I]$ is a σ-algebra] hence $U \in \sigma[\mathcal{I}]$

so $\mathcal{T}_{\mathbb{R}} \subseteq \sigma[\mathcal{I}]$ and thus

$$\sigma[\mathcal{T}_{\mathbb{R}}] \subseteq \sigma[\mathcal{I}] \tag{18.98}$$

Further if $I \in \mathcal{I}$ then we have either

$\boldsymbol{I = \emptyset}$**.** then $I \in \sigma[\mathcal{T}_{\mathbb{R}}]$

$\boldsymbol{I \neq \emptyset}$**.** then using 18.56 there exists a $\{U_i\}_{i \in \mathbb{N}} \subseteq \mathcal{T}_{\mathbb{R}}$ such that $I = \bigcap_{i \in \mathbb{N}} I_i \in \sigma[\mathcal{T}_{\mathbb{R}}]$ [because $\sigma[\mathcal{T}_{\mathbb{R}}]$ is a σ-algebra together with 18.8]

so $\mathcal{I} \in \sigma[\mathcal{T}_{\mathbb{R}}]$ and thus

$$\sigma[\mathcal{I}] \subseteq \sigma[\mathcal{T}_{\mathbb{R}}] \tag{18.99}$$

The proposition is then proved by 18.98 and 18.99 □

Actually we can also generate $\sigma[\mathcal{I}]$ by closed intervals as shown in the following proposition.

Proposition 18.58. *We have* $\sigma[\{[a,b] | a, b \in \mathbb{R} \,\text{with}\, a \leqslant b\}] = \mathcal{B}[\mathbb{R}]$

Proof. Take $a, b \in \mathbb{R}$ with $a \leqslant b$ then

$$\begin{aligned}
x \in [a, b+1[\setminus]b, b+1[&\Leftrightarrow a \leqslant x < b+1 \wedge \neg(b < x < b+1) \\
&\Leftrightarrow a \leqslant x < b+1 \wedge (x \leqslant b \vee b+1 \leqslant x) \\
&\Leftrightarrow (a \leqslant x < b+1 \wedge x \leqslant b) \vee (a \leqslant x < b+1 \wedge b+1 \leqslant x) \\
&\Leftrightarrow a \leqslant x < b+1 \wedge x \leqslant b \\
&\Leftrightarrow a \leqslant x \leqslant b \\
&\Leftrightarrow x \in [a, b]
\end{aligned}$$

proving that

$$[a, b] = [a, b+1[\setminus]b, b+1[\tag{18.100}$$

Now $[a, b+1[\in \mathcal{I} \subseteq \mathcal{B}[\mathbb{R}]$ (see 18.57) and $]b, b+1[\in \mathcal{T}_{\mathbb{R}} \subseteq \mathcal{B}[\mathbb{R}]$ hence using the properties of of a σ-algebra (see 18.8) we have that $[a, b+1[\setminus]b, b+1[\in \mathcal{B}[\mathbb{R}]$. So $\{[a,b] | a, b \in \mathbb{R} \,\text{with}\, a \leqslant b\} \subseteq \mathcal{B}[\mathbb{R}]$ and as $\sigma[\{[a,b] | a, b \in \mathbb{R} \,\text{with}\, a \leqslant b\}]$ is the smallest σ-algebra containing $\{[a,b] | a, b \in \mathbb{R} \,\text{with}\, a \leqslant b\}$ we have

$$\sigma[\{[a,b] | a, b \in \mathbb{R} \,\text{with}\, a \leqslant b\}] \subseteq \mathcal{B}[\mathbb{R}] \tag{18.101}$$

For the opposite inclusion note let $I \in \mathcal{I}$ then we have either

$\boldsymbol{I = \emptyset}$**.** then $I = \emptyset \in \sigma[\{[a,b] | a, b \in \mathbb{R} \,\text{with}\, a \leqslant b\}]$

$\boldsymbol{I \neq \emptyset}$**.** then there exists a $a, b \in \mathbb{R}$ with $a \leqslant b$. Note that

$$\begin{aligned}
x \in [a, b+1] \setminus [b, b+1] &\Leftrightarrow a \leqslant x \leqslant b+1 \wedge \neg(b \leqslant x \leqslant b+1) \\
&\Leftrightarrow a \leqslant x \leqslant b+1 \wedge (x < b \vee b+1 < x) \\
&\Leftrightarrow (a \leqslant x \leqslant b \wedge x < b) \vee (a \leqslant x \leqslant b \wedge b+1 < x) \\
&\Leftrightarrow (a \leqslant x \leqslant b \wedge x < b) \\
&\Leftrightarrow a \leqslant x < b \\
&\Leftrightarrow x \in [a, b[
\end{aligned}$$

proving that $[a,b[=[a,b+1]\setminus[b,b+1]\in\sigma[\{[a,b]|a,b\in\mathbb{R}\,\text{with}\,a\leqslant b\}]$ [see 18.8]

So we have that $\mathcal{I}\subseteq\sigma[\{[a,b]|a,b\in\mathbb{R}\text{ with }a\leqslant b\}]$ hence $\mathcal{B}[\mathbb{R}]=\sigma[\mathcal{I}]\subseteq\sigma[\{[a,b]|a,$ $b\in\mathbb{R}\text{ with }a\leqslant b\}]$ proving together with 18.101 that

$$[\{[a,b]|a,b\in\mathbb{R}\,\text{with}\,a\leqslant b\}]=\mathcal{B}[\mathbb{R}] \qquad \square$$

Lemma 18.59. *We have* $\sigma[\{]-\infty,a]|a\in\mathbb{R}\}]=\sigma[\{]-\infty,a[|a\in\mathbb{R}\}]$

Proof. Let $a\in\mathbb{R}$ then we have $\forall n\in\mathbb{N}$ that $]-\infty,a]\subseteq\,]-\infty,a+\frac{1}{n}[$ so that $]-\infty,$ $a]\subseteq\bigcap_{n\in\mathbb{N}}]-\infty,a+\frac{1}{n}[$. For the opposite equality take $x\in\bigcap_{n\in\mathbb{N}}]-\infty,a+\frac{1}{n}[$ then $\forall n\in\mathbb{N}$ we have $x<a+\frac{1}{n}$ so using 9.56 (3) we have $x\leqslant a$ proving that $\bigcap_{n\in\mathbb{N}}]-\infty,$ $a+\frac{1}{n}[\,\subseteq]-\infty,a]$ giving $]-\infty,a]=\bigcap_{n\in\mathbb{N}}]-\infty,a+\frac{1}{n}[\,\in\sigma[\{]-\infty,a[|a\in\mathbb{R}\}]$ [because of 18.8]. Hence we have $\{]-\infty,a]|a\in\mathbb{R}\}\subseteq\sigma[\{]-\infty,a[|a\in\mathbb{R}\}]$ giving

$$\sigma[\{]-\infty,a]|a\in\mathbb{R}\}]\subseteq\sigma[\{]-\infty,a[|a\in\mathbb{R}\}] \tag{18.102}$$

For the opposite inclusion note that given $a\in\mathbb{R}$, $\forall n\in\mathbb{N}$ we have $]-\infty,a-\frac{1}{n}]\subseteq$ $]-\infty,a[$ proving that $\bigcup_{n\in\mathbb{N}}]-\infty,a-\frac{1}{n}]\subseteq]-\infty,a[$. Take $x\in]-\infty,a[$ then $x<a$ or $0<a-x$, using 9.55 (3) there exists a $n\in\mathbb{N}$ such that $0<\frac{1}{n}<a-x\Rightarrow x<a-\frac{1}{n}\Rightarrow$ $x\in\,]-\infty,a-\frac{1}{n}[$ proving that $]-\infty,a[\,\subseteq\bigcup_{n\in\mathbb{N}}]-\infty,a-\frac{1}{n}]$, giving $]-\infty,a[=$ $\bigcup_{n\in\mathbb{N}}]-\infty,a-\frac{1}{n}]\in\sigma[]-\infty,a]|a\in\mathbb{R}]$. Hence we have $\sigma[\{]-\infty,a[|a\in\mathbb{R}\}]\subseteq\sigma[]-\infty,$ $a]|a\in\mathbb{R}]$ which using (18.102) means that

$$\sigma[\{]-\infty,a]|a\in\mathbb{R}\}]=\sigma[\{]-\infty,a[|a\in\mathbb{R}\}] \qquad \square$$

Lemma 18.60. *We have* $\sigma[\{]a,\infty[|a\in\mathbb{R}\}]=\sigma[\{]-\infty,a]|a\in\mathbb{R}\}]$

Proof. Let $a\in\mathbb{R}$ then

$$\begin{aligned} x\in\mathbb{R}\setminus]-\infty,a] &\Leftrightarrow \neg(x\leqslant a)\\ &\Leftrightarrow a<x\\ &\Leftrightarrow x\in]a,\infty[\end{aligned}$$

proving that $]a,\infty[=\mathbb{R}\setminus]-\infty,a]\in\sigma[\{]-\infty,a]|a\in\mathbb{R}\}]$. So $\{]a,\infty[|a\in\mathbb{R}\}\subseteq\sigma[\{]-\infty,$ $a]|a\in\mathbb{R}\}]$ hence

$$\sigma[\{]a,\infty[|a\in\mathbb{R}\}]\subseteq\sigma[\{]-\infty,a]|a\in\mathbb{R}\}] \tag{18.103}$$

Further if $a\in\mathbb{R}$ then

$$\begin{aligned} x\in\mathbb{R}\setminus]a,\infty[&\Leftrightarrow \neg(a<x)\\ &\Leftrightarrow x\leqslant a\\ &\Leftrightarrow]-\infty,a]\end{aligned}$$

proving that $]-\infty, a] = \mathbb{R} \setminus]a, \infty[\in \sigma[\{]a, \infty[| a \in \mathbb{R}\}]$. So $\{]-\infty, a] | a \in \mathbb{R}\} \subseteq \sigma[\{]a, \infty[| a \in \mathbb{R}\}]$, hence $\sigma[\{]-\infty, a] | a \in \mathbb{R}\}] \subseteq \sigma[\{]a, \infty[| a \in \mathbb{R}\}]$ which together with (18.103) proves that

$$\sigma[\{]a, \infty[| a \in \mathbb{R}\}] = \sigma[\{]-\infty, a] | a \in \mathbb{R}\}] \qquad \square$$

Lemma 18.61. *We have* $\sigma[\{[a, \infty[| a \in \mathbb{R}\}] = \sigma[\{]-\infty, a[| a \in \mathbb{R}\}]$

Proof. Let $a \in \mathbb{R}$ then

$$\begin{aligned} x \in \mathbb{R} \setminus]-\infty, a[&\Leftrightarrow \neg(x < a) \\ &\Leftrightarrow a \leqslant x \\ &\Leftrightarrow x \in [a, \infty[\end{aligned}$$

proving that $[a, \infty[= \mathbb{R} \setminus]-\infty, a[\in \sigma[\{]-\infty, a[| a \in \mathbb{R}\}]$. Hence $\{[a, \infty[| a \in \mathbb{R}\} \subseteq \sigma[\{]-\infty, a[| a \in \mathbb{R}\}]$ giving

$$\sigma[\{[a, \infty[| a \in \mathbb{R}\}] \subseteq \sigma[\{]-\infty, a[| a \in \mathbb{R}\}] \tag{18.104}$$

Further given $a \in \mathbb{R}$ we have

$$\begin{aligned} x \in \mathbb{R} \setminus [a, \infty[&\Leftrightarrow \neg(a \leqslant x) \\ &\Leftrightarrow x < a \\ &\Leftrightarrow x \in]-\infty, a[\end{aligned}$$

proving that $]-\infty, a[= \mathbb{R} \setminus [a, \infty[\in \sigma[\{[a, \infty[| a \in \mathbb{R}\}]$. Hence $\{]-\infty, a[| a \in \mathbb{R}\} \subseteq \sigma[\{[a, \infty[| a \in \mathbb{R}\}]$ giving $\sigma[\{]-\infty, a[| a \in \mathbb{R}\}] \subseteq \sigma[\{[a, \infty[| a \in \mathbb{R}\}]$. Together with (18.104) it follows that

$$\sigma[\{]-\infty, a[| a \in \mathbb{R}\}] = \sigma[\{[a, \infty[| a \in \mathbb{R}\}] \qquad \square$$

To summarize the above lemmas we have by 18.48, 18.49, 18.57,18.58,18.59,18.60 and 18.61 that

Theorem 18.62. *For $\mathcal{L}$ the set of Lebesgue measurable sets of $\mathbb{R}$ (see Lebesgue measurable set in 18.45) we have*

$$\begin{aligned} \mathcal{B}[\mathbb{R}] &= \sigma[\{[a, b[| a, b \in \mathbb{R} \text{ with } a \leqslant b\}] \\ &= \sigma[\{]a, b[| a, b \in \mathbb{R} \text{ with } a \leqslant b\}] \\ &= \sigma[\{[a, b] | a, b \in \mathbb{R} \text{ with } a \leqslant b\}] \\ &= \sigma[\{[a, \infty[| a \in \mathbb{R}\}] \\ &= \sigma[\{]a, \infty[| a \in \mathbb{R}\}] \\ &= \sigma[\{]-\infty, a] | a \in \mathbb{R}\}] \\ &= \sigma[\{]-\infty, a[| a \in \mathbb{R}\}] \\ &\subseteq \mathcal{L} \end{aligned}$$

constructing the measure space $\langle \mathbb{R}, \mathcal{B}[\mathbb{R}], \lambda_{|\mathcal{B}[\mathbb{R}]} \rangle$

18.1.4.2 Lebesgue measure on $\mathbb{R}^n$

We will develop our theory of a Lebesgue measure on $\mathbb{R}^n$ similar to the Lebesgue measure on $\mathbb{R}^n$ so the first thing we do is define the concept of half open intervals in $\mathbb{R}^n$.

Definition 18.63. *Let $a, b \in \mathbb{R}^n$ then we say that*

1. *$a \leqslant b$ if $\forall i \in \{1, \ldots, n\}$ we have $a_i \leqslant b_i$*
2. *$a < b$ if $\forall i \in \{1, \ldots, n\}$ we have $a_i < b_i$*

Definition 18.64. *Let $n \in \mathbb{N}$ and $a, b \in \mathbb{R}^n$ with $a \leqslant b$ then*

$$\begin{aligned}
[a, b[&= \prod_{i \in \{1, \ldots, n\}} [a_i, b_i[\\
&= \{x \in \mathbb{R}^n | \forall i \in \{1, \ldots, n\} \text{ we have } x_i \in [a_i, b_i[\} \\
&= \{x \in \mathbb{R}^n | \forall i \in \{1, \ldots, n\} \text{ we have } a_i \leqslant x_i < b_i\} \\
[a, b] &= \prod_{i \in \{1, \ldots, n\}} [a_i, b_i] \\
&= \{x \in \mathbb{R}^n | \forall i \in \{1, \ldots, n\} \text{ we have } x_i \in [a_i, b_i]\} \\
&= \{x \in \mathbb{R}^n | \forall i \in \{1, \ldots, n\} \text{ we have } a_i \leqslant x_i \leqslant b_i\} \\
]-\infty, a[&= \prod_{i \in \{1, \ldots, n\}}]-\infty, a_i[\\
&= \{x \in \mathbb{R}^n | \forall i \in \{1, \ldots, n\} \text{ we have } x_i \in]-\infty, a_i[\} \\
&= \{x \in \mathbb{R}^n | \forall i \in \{1, \ldots, n\} \text{ we have } x_i < a_i\}
\end{aligned}$$

the set of all half-open intervals in $\mathbb{R}^n$ is called $\mathcal{R}^n$, so $\mathcal{R}^n = \{[a, b[| a, b \in \mathbb{R}^n \text{ with } a \leqslant b\}$

The condition $a \leqslant b$ in the above definition is there to allow the empty set in the set of all half-open intervals in $\mathbb{R}^n$ as is expressed in the following proposition.

Proposition 18.65. *Let $n \in \mathbb{R}^n$ then $[a, b[\in \mathcal{R}^n$ is empty if and only if $\exists i \in \{1, \ldots, n\}$ such that $a_i = b_i$*

Proof. Let $[a, b[= \emptyset$ and assume that $\forall i \in \{1, \ldots, n\}$ $a_i < b_i$ then $a = (a_1, \ldots, a_n) \in [a, b[$ contradicting $[a, b[= \emptyset$. Hence if $[a, b[= \emptyset$ then there exists a $i \in \{1, \ldots, n\}$ such that $a_i = b_i$. On the other hand if there exists a $i \in \{1, \ldots, n\}$ such that $a_i = b_i$ then if $x \in [a, b[$ we have $a_i \leqslant x < a_i$ a contradiction so $[a, b[= \emptyset$ □

The endpoints of a non empty half-open intervals in $\mathbb{R}^n$ are unique just like endpoints in $\mathbb{R}$.

Proposition 18.66. *Let $n \in \mathbb{N}$ and $\emptyset \neq R \in \mathcal{R}^n$ then there exists a unique $a, b \in \mathbb{R}^n$ such that $R = [a, b[$.*

Proof. By the definition of $\mathcal{R}^n$ there exists a $a, b \in \mathbb{R}^n$ with $a \leqslant b$ such that $R = [a, b[$ which proves existence. Assume that also $R = [a', b'[$ Using 5.103 and the fact that $R \neq \emptyset$ we must have then that $\forall i \in \{1, \ldots, n\}$ we have that $[a_i, b_i[= [a'_i, b'_i[\neq \emptyset$ hence using 18.39 we must have that $a_i = a'_i$ and $b_i = b'_i$ proving that $a = a' \wedge b = b'$. □

Using the above we can define then the volume of a interval in $\mathbb{R}^n$ and begin or end of a interval in $\mathbb{R}^n$

Definition 18.67. *Let $n \in \mathbb{N}$ then* $\text{begin}: \mathcal{R}^n \setminus \{\emptyset\} \rightarrow \mathbb{R}^n$ *and* $\text{end}: \mathcal{R}^n \setminus \{\emptyset\} \rightarrow \mathbb{R}^n$ *are defined by* $\text{begin}(R) = a$ *and* $\text{end}(R) = b$ *where* $R = [a, b[$.

Definition 18.68. *Let $n \in \mathbb{N}$ then $v^n: \mathcal{R} \rightarrow [0, \infty]$ is defined by $v^n(\emptyset) = 0$ and if $R \neq \emptyset$ $v^n(R) = \prod_{i=1}^n l([a_i, b_i[)$ where $R = [a, b[= \prod_{i \in \{1, \ldots, n\}} [a_i, b_i[$*

Note 18.69. If $n = 1$ then $\mathbb{R}^n = 1$ and if $R \in \mathcal{R}^1$ then there exists a $a, b \in \mathbb{R}$ with $a_1 = a \leqslant b = b_1$ such that $R = [a, b[= \prod_{i \in \{1\}} [a_i, b_i[$ so that $\mathcal{R}^1 = \mathcal{I}$. Further we have that $v^1(R) = \prod_{i=1}^1 l([a_i, b_i[) = l(R)$

Lemma 18.70. *Let $n \in \mathbb{N}$, $R \in \mathcal{R}^n$ and $\{R_i\}_{i \in \mathbb{N}} \subseteq \mathcal{R}^n$ such that $R \subseteq \bigcup_{i \in \mathbb{N}} R_i$ then $v^n(R) \leqslant \sum_{i=1}^\infty v^n(R_i)$*

Proof. we prove this by induction on n so let $\mathcal{S} = \{n \in \mathbb{N} | \forall R \in \mathcal{R}^n, \forall \{R_i\}_{i \in \mathbb{N}} \subseteq \mathcal{R}^n \text{ with } R \subseteq \bigcup_{i \in \mathbb{N}} R_i \text{ we have } v^n(R) \leqslant \sum_{i=1}^\infty v^n(R_i)\}$ then

$1 \in \mathcal{S}$. If $R \in \mathcal{R}^1$ and $\{R_i\}_{i \in \mathbb{N}} \in \mathcal{R}^1$ then (see 18.69) we have that $R \in \mathcal{I}$ and $\{R_i\}_{i \in \mathbb{N}} \subseteq \mathcal{I}$. So we can apply 18.43, giving $v^1(R) = l(R) \leqslant \sum_{i=1}^\infty l(R_i) = \sum_{i=1}^\infty v^n(R_i)$, proving that $1 \in \mathcal{S}$

$n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$. Let $R \in \mathcal{R}^{n+1}$ and $\{R_i\}_{i \in \mathbb{N}} \subseteq \mathcal{R}^{n+1}$ such that $R \subseteq \bigcup_{i \in \mathbb{N}} R_i$, then $R = [a, b[$ and $R_i = [a^i, b^i[$ where $a, b, a^i, b^i \in \mathcal{R}^{n+1}$ such that $a \leqslant b$ and $a_i \leqslant b_i$. If we write

$$V = \prod_{i=1}^n (b_i - a_i) \geqslant 0 \tag{18.105}$$

we have

$$\begin{aligned} v^{n+1}(R) &= v^{n+1}([a, b[) \\ &= v^{n+1}\left(\prod_{i \in \{1, \ldots, n+1\}} [a_i, b_i[\right) \\ &= \prod_{i=1}^{n+1} (b_i - a_i) \\ &= \left(\prod_{i=1}^n (b_i - a_i) \right) \cdot (b_{n+1} - a_{n+1}) \\ &= V \cdot (b_{n+1} - a_{n+1}) \end{aligned}$$

proving

$$v^{n+1}(R) = V \cdot (b_{n+1} - a_{n+1}) \tag{18.106}$$

Take now $\varepsilon > 0$ then for $r \in \mathbb{R}$ with $a_{n+1} \leqslant r$ define $H_r = \{x \in \mathbb{R}^{n+1} | x_{n+1} < r\}$. Then given $i \in \mathbb{N}$ $\forall x \in R_i \bigcap H_r$ we have $x_j \in [a_j^i, b_j^i[\Rightarrow a_j^i \leqslant x_j < b_j^i$ $(j \in \{1, \ldots, n+1\})$ and $x_{n+1} < r$ which is equivalent to $x \in [a^i, \bar{b}^i[$ where $\bar{b}_j^i = \begin{cases} b_j^i \text{ if } j \in \{1, \ldots, n\} \\ \min(b_{n+1}^i, r) \end{cases}$ (so that $a \leqslant \bar{b}$ and thus that $R_i = [a, \bar{b}[)$ proving that $v^{n+1}(R_i \bigcap H_r)$ is defined and

$$
\begin{aligned}
v^{n+1}\left(R_i \bigcap H_r\right) &= v^{n+1}\left(\prod_{j \in \{1, \ldots, n+1\}} [a_j^i, \bar{b}_j^i[\right) \\
&= \prod_{j=1}^{n+1} (\bar{b}_j^i - a_j^i) \\
&= \left(\prod_{j=1}^{n} (\bar{b}_j^i - a_j^i)\right) \cdot (\bar{b}_{n+1}^i - a_{n+1}^i) \\
&= \left(\prod_{j=1}^{n} (b_j^i - a_j^i)\right) \cdot (\min(b_{n+1}^i, r) - a_j^i)
\end{aligned} \tag{18.107}
$$

Further we have from the above also

$$\forall r, s \text{ with } a_{n+1} \leqslant r \leqslant s \text{ that } v^{n+1}\left(R_i \bigcap H_r\right) \leqslant v^{n+1}\left(R_i \bigcap H_s\right) \forall i \in \mathbb{N} \tag{18.108}$$

So the following definition is valid

$$A_\varepsilon = \left\{x \in | a_{n+1} \leqslant x \leqslant b_{n+1} \wedge V \cdot (x - a_{n+1}) \leqslant (1 + \varepsilon) \cdot \sum_{i=1}^{\infty} v^{n+1}\left(R_i \bigcap H_x\right)\right\} \tag{18.109}$$

using the above definition we have

$$A_\varepsilon \subseteq [a_{n+1}, b_{n+1}] \tag{18.110}$$

As $V \cdot (a_{n+1} - a_{n+1}) = 0 \leqslant (1 + \varepsilon) \cdot \sum_{i=1}^{\infty} v^{n+1}(R_i \bigcap H_x)$ and $a_{n+1} \leqslant a_{n+1} \leqslant b_{n+1}$ we have that $a_{n+1} \in A_\varepsilon$ proving that $A_\varepsilon \neq \emptyset$, as also by 18.110 b_{n+1} is a upper bound of A_ε, we have by the conditional completeness of the real numbers (see 9.43) that

$$\gamma = \sup(A_\varepsilon) \text{ exists and } a_{n+1} \leqslant \gamma \leqslant b_{n+1} \text{ (see (18.110))} \tag{18.111}$$

Further as for $x \in A_\varepsilon$ we have $a_{n+1} \leqslant x \leqslant \gamma$ we have by 18.108 that $v^{n+1}(R_i \bigcap H_x) \leqslant v^{n+1}(R_i \bigcap H_\gamma)$ proving that

$$\forall x \in A_\varepsilon \text{ we have } V \cdot (x - a_{n+1}) \leqslant (1 + \varepsilon) \cdot \sum_{i=1}^{\infty} v^{n+1}\left(R_i \bigcap H_x\right) \leqslant (1 + \varepsilon) \cdot \sum_{i=1}^{\infty} v^{n+1}\left(R_i \bigcap H_\gamma\right) \tag{18.112}$$

Now

$$\begin{aligned} V\cdot(\gamma-a_{n+1}) &= V\cdot(\sup(A_\varepsilon)-a_{n+1}) \\ &\underset{17.29}{=} V\cdot\sup(\{x-a_{n+1}|x\in A_\varepsilon\}) \\ &\underset{0\leqslant V \text{ and Algorithm } 17.30}{=} \sup(\{V\cdot(x-a_{n+1})|x\in A_\varepsilon\}) \\ &\leqslant_{(18.112)} (1+\varepsilon)\cdot\sum_{i=1}^{\infty} v^{n+1}\left(R_i\bigcap H_\gamma\right) \end{aligned} \tag{18.113}$$

which proves together with (18.111) that

$$\gamma\in A_\varepsilon \tag{18.114}$$

Assume now that $\gamma<b_{n+1}$. Using 18.111 we have then that

$$\gamma\in[a_{n+1},b_{n+1}[\tag{18.115}$$

Set

$$J=\{x\in\mathbb{R}^n|(x,\gamma)\in R\} \text{ and } J_j=\{x\in\mathbb{R}^n|(x,\gamma)\in R_j\}\ j\in\mathbb{N} \tag{18.116}$$

Now if we take $a'=(a_1,\ldots,.a_n), b'=(b_1,\ldots,b_n)\in\mathbb{R}^n$ then as $a\leqslant b$ we have $a'\leqslant b'$ and using (18.115) we have that

$$\begin{aligned} x\in[a',b'[&\underset{(18.115)}{\Leftrightarrow} a'\leqslant x<b'\wedge a_{n+1}\leqslant\gamma<b_{n+1} \\ &\Leftrightarrow \forall i\in\{1,\ldots,n\}\ a_i\leqslant x_i<b_i\wedge a_{n+1}\leqslant\gamma<b_{n+1} \\ &\Leftrightarrow (x,\gamma)\in R \\ &\Leftrightarrow x\in J \end{aligned}$$

proving that

$$J=[a',b'[\in\mathcal{R}^n \tag{18.117}$$

Given $j\in\mathbb{N}$ we have either

$\boldsymbol{\gamma\in[a^j_{n+1},b^j_{n+1}[}$**.** then if we take $a'^j=(a^j_1,\ldots,a^j_n), b'^j=(b^j_1,\ldots,b^j_n)\in\mathbb{R}^n$ we have from $a^j\leqslant b^j$ that $a'^j\leqslant b'^j$. Further

$$\begin{aligned} x\in J_j &\Leftrightarrow (x,\gamma)\in R_j \\ &\Leftrightarrow \forall i\in\{1,\ldots,n\}\ a^j_i\leqslant x_i<b^j_i \text{ and } a^j_{n+1}\leqslant\gamma<b^j_{n+1} \\ &\underset{\gamma\in[a^j_{n+1},b^j_{n+1}[}{\Leftrightarrow} \forall i\in\{1,\ldots,n\}\ a^j_i\leqslant x_i<b^j_i \\ &\Leftrightarrow x\in[a'^j,b'^j[\in\mathcal{R}^n \end{aligned}$$

$\boldsymbol{\gamma\notin[a^j_{n+1},b^j_{n+1}[}$**.** then $J_i=\emptyset\in\mathcal{R}^n$

so in all cases we have

$$J_j\in\mathcal{R}^n \text{ and either } a^j_{n+1}\leqslant\gamma<b^j_{n+1}\Rightarrow J_j=[a'^j,b'^j[\Rightarrow v^n(J_j)=\prod_{i=1}^{n}(b^j_i-a^j_i) \text{ or } J_j=\emptyset \tag{18.118}$$

Now if $x \in J$ then $(x,\gamma) \in R \subseteq \bigcup_{i \in \mathbb{N}} R_i$ hence there exists a $i \in \mathbb{N}$ such that $(x,\gamma) \in R_i$ proving that $x \in J_i$. So

$$J \subseteq \bigcup_{i \in \mathbb{N}} J_i \tag{18.119}$$

As $n \in \mathcal{S}$ we have combining 18.117, (18.118) and (18.119) we have $v^n(J) \leqslant \sum_{i=1}^{\infty} v^n(J_i)$, further $v^n(J) = \prod_{i=1}^{n} (b_i' - a_i') = \prod_{i=1}^{n} (b_i - a_i) \underset{(18.105)}{=} V$ so we have

$$V = v^n(J) \leqslant \sum_{i=1}^{\infty} v^n(J_i) \tag{18.120}$$

As $\emptyset \neq R = \prod_{i=1}^{n+1} [a_i, b_i[$ we have by 5.102 that $\forall i \in \{1, \ldots, n+1\}$ $[a_i, b_i[\neq \emptyset$ hence $a_i < b_i$ proving that $0 < V$. As we have taken $\varepsilon > 0$ we have that $1 + \varepsilon > 1$ so that $\frac{V}{1+\varepsilon} < V \leqslant \sum_{i=1}^{\infty} v^n(J_i) \underset{17.113}{=} \sup(\{\sum_{i=1}^{m} v^n(J_i) | m \in \mathbb{N}\})$. Using the definition of a supremum it follows then that there exists a $m \in \mathbb{N}$ such that $\frac{V}{1+\varepsilon} < \sum_{i=1}^{m} v^n(J_i) \leqslant \sum_{i=1}^{\infty} v^n(J_i)$. So

$$V < (1+\varepsilon) \cdot \sum_{i=1}^{m} v^n(J_i) \tag{18.121}$$

Define

$$\chi = \min\left(\{b_{n+1}\} \bigcup \{b_{n+1}^j | j \in \{j \in \{1, \ldots, m\} | J_j \neq \emptyset\}\}\right) \tag{18.122}$$

Now as by 18.118 we have $\forall j \in \{1, \ldots, m\}$ with $J_j \neq \emptyset$ that $\gamma < b_{n+1}^j$ and we assumed that $\gamma < b_{n+1}$ we have that

$$b_{n+1} \geqslant \mathcal{X} > \gamma \geqslant a_{n+1} \tag{18.123}$$

so that $v^n(R_j \bigcap H_{\mathcal{X}})$ and $v^n(R_j \bigcap H_\gamma)$ is defined. For $j \in \{1, \ldots, m\}$ we have either (see (18.118))

$J_j = \emptyset$. then

$$\begin{aligned} v^{n+1}\left(R_j \bigcap H_\gamma\right) + (\mathcal{X} - \gamma) \cdot v^n(J_j) &= v^{n+1}\left(R_j \bigcap H_\gamma\right) + 0 \\ &\leqslant_{(18.123) \text{ and } (18.108)} v^{n+1}\left(R_j \bigcap H_\chi\right) \end{aligned}$$

$J_j \neq \emptyset$. then $a_{n+1}^j \leqslant \gamma < b_{n+1}^j$ and

$$\begin{aligned} v^{n+1}\left(R_j \bigcap H_\chi\right) &\underset{(18.107)}{=} \left(\prod_{i=1}^{n} (b_i^j - a_i^j)\right) \cdot (\min(b_{n+1}^j, \mathcal{X}) - a_{n+1}^j) \\ &\underset{(18.123)}{=} \left(\prod_{i=1}^{n} (b_i^j - a_i^j)\right) \cdot (\chi - a_{n+1}^j) \\ &= \left(\prod_{i=1}^{n} (b_i^j - a_i^j)\right) \cdot (\chi - \gamma + \gamma - a_{n+1}^j) \end{aligned}$$

$$
\begin{aligned}
&= \left(\prod_{i=1}^{n}(b_i^j - a_i^j)\right)\cdot(\mathcal{X}-\gamma) + \left(\prod_{i=1}^{n}(b_i^j - a_i^j)\right)\cdot (\gamma - a_{n+1}^j) \\
&\underset{(18.118)}{=} v^n(J_j)\cdot(\chi-\gamma) + \left(\prod_{i=1}^{n}(b_i^j - a_i^j)\right)\cdot(\gamma - a_{n+1}^j) \\
&= v^n(J_j)\cdot(\chi-\gamma) + \left(\prod_{i=1}^{n}(b_i^j - a_i^j)\right)\cdot (\min(b_{n+1}^j,\gamma) - a_{n+1}^j) \\
&\underset{(18.107)}{=} v^n(J_j)\cdot(\chi-\gamma) + v^{n+1}\big(R_j \bigcap H_\gamma\big)
\end{aligned}
$$

so we have in all cases

$$\forall j\in\{1,\ldots,m\} \text{ we have } v^{n+1}\big(R_j\bigcap H_\gamma\big) + (\chi-\gamma)\cdot v^n(J_j) \leqslant v^{n+1}\big(R_j \bigcap H_\chi\big) \tag{18.124}$$

Further

$$
\begin{aligned}
V\cdot(\chi - a_{n+1}) &= V\cdot(\chi-\gamma+\gamma-a_{n+1}) \\
&= V\cdot(\gamma-a_{n+1}) + V\cdot(\chi-\gamma) \\
&\leqslant_{(18.113)} (1+\varepsilon)\cdot\sum_{i=1}^{\infty} v^{n+1}\big(R_i\bigcap H_\gamma\big) + V\cdot(\chi-\gamma) \\
&\leqslant_{(18.121)} (1+\varepsilon)\cdot\sum_{i=1}^{\infty} v^{n+1}\big(R_i\bigcap H_\gamma\big) + (\chi-\gamma)\cdot (1+\varepsilon)\cdot\sum_{i=1}^{m} v^n(J_i) \\
&= (1+\varepsilon)\cdot\sum_{i=m+1}^{\infty} v^{n+1}\big(R_i\bigcap H_\gamma\big) + (1+\varepsilon)\cdot\sum_{i=1}^{m} v^{n+1}\big(R_i\bigcap H_\gamma\big) + (\chi-\gamma)\cdot(1+\varepsilon)\cdot\sum_{i=1}^{m} v^n(J_i) \\
&= (1+\varepsilon)\cdot\sum_{i=m+1}^{\infty} v^{n+1}\big(R_i\bigcap H_\gamma\big) + (1+\varepsilon)\cdot\sum_{i=1}^{m}\big(v^{n+1}\big(R_i\bigcap H_\gamma\big) + (\chi-\gamma)\cdot v^n(J_i)\big)
\end{aligned}
$$

$$
\begin{aligned}
&\leqslant_{(18.124)} && (1+\varepsilon)\cdot\sum_{i=m+1}^{\infty} v^{n+1}\Big(R_i\bigcap H_\gamma\Big)+(1+\\
& && \varepsilon)\cdot\sum_{i=1}^{m} v^{n+1}\Big(R_i\bigcap H_\chi\Big)\\
&\leqslant_{(18.123)\text{ and }(18.108)} && (1+\varepsilon)\cdot\sum_{i=m+1}^{\infty} v^{n+1}\Big(R_i\bigcap H_\chi\Big)+(1+\\
& && \varepsilon)\cdot\sum_{i=1}^{m} v^{n+1}\Big(R_i\bigcap H_\chi\Big)\\
&= && (1+\varepsilon)\cdot\sum_{i=1}^{\infty} v^{n+1}\Big(R_i\bigcap H_\chi\Big)
\end{aligned}
$$

proving using 18.123 that $\chi\in A_\varepsilon$ so that $\chi\leqslant\sup(A_\varepsilon)=\gamma$ contradicting (18.123). Hence the assumption $\gamma<b_{n+1}$ must be wrong. Combining this with 18.111 proves that $b_{n+1}=\gamma\in A_\varepsilon$ (see (18.114)). So by the definition of A_ε we have that

$$
\begin{aligned}
V\cdot(b_{n+1}-a_{n+1}) &\leqslant && (1+\varepsilon)\cdot\sum_{i=1}^{\infty} v^{n+1}\Big(R_i\bigcap H_{b_{n+1}}\Big)\\
&\underset{(18.107)}{=} && (1+\varepsilon)\cdot\sum_{i=1}^{\infty}\left[\left(\prod_{j=1}^{n}(b_j^i-a_j^i)\right)\cdot(\min(b_{n+1}^i,b_{n+1}^i)-\right.\\
& && \left.a_{n+1}^i)\right]\\
&= && (1+\varepsilon)\cdot\sum_{i=1}^{\infty}\left[\left(\prod_{j=1}^{n}(b_i^i-a_j^i)\right)\cdot(b_{n+1}^i-a_{n+1}^i)\right]\\
&\underset{(18.107)}{=} && (1+\varepsilon)\cdot\sum_{i=1}^{\infty} v^{n+1}(R_i) \qquad (18.125)
\end{aligned}
$$

As we have chosen ε arbitrary it follows (using 9.56 (6)) that

$$
v^{n+1}(R)\underset{(18.106)}{=}V\cdot(b_{n+1}-a_{n+1})\leqslant\sum_{i=1}^{\infty} v^{n+1}(R_i) \qquad \square
$$

We can now define a outer measure based on the volume of a half open interval.

Theorem 18.71. *Let $n\in\mathbb{N}$ then $\varphi^n\colon\mathcal{P}(\mathbb{R}^n)\to[0,\infty]$ defined by $\varphi^n(A)=\inf(\{\sum_{i=1}^{\infty}v^n(A_i)|\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{R}^n\text{ such that }A\subseteq\bigcup_{i\in\mathbb{N}}A_i\})$ then φ^n is a outer-measure. Further $\forall A\in\mathcal{R}^n$ we have $v^n(A)=\varphi^n(A)$.*

Proof. *We have*

1. $\emptyset = \prod_{i \in \{1,\ldots,n\}} [i, i[$

2. *Let* $x \in \mathbb{R}^n$ *then using a consequence of the Archimedean property of the real numbers (see 9.55) there exists* $\forall i \in \{1, \ldots, n\}$ *a* $n_i \in \mathbb{N}$*, such that* $|x_i| < n_i \Rightarrow x_i, -x_i < n_i \Rightarrow x_i \in [-n_i, n_i[$*, proving that* $x \in \prod_{i \in \{1,\ldots,n\}} [-n_i, n_i[$*. Take* $n_x = \max(\{n_i | i \in \{1, \ldots, n\}\}) \in \mathbb{N}$ *then* $x \in \prod_{i \in \{1,\ldots,n\}} [-n_x, n_x[$*. Hence if we take* $\{R_j\}_{j \in \mathbb{N}}$ *by* $R_j = \prod_{i \in \{1,\ldots,n\}} [-j, j[\subseteq \mathcal{R}^n$ *we have*

$$\mathbb{R}^n \subseteq \bigcup_{j \in \mathbb{N}} R_j$$

3. $\varphi^n(\emptyset) = 0$ *by definition*

4. *Using 18.70 we have* $\forall A \in \mathcal{R}^n, \{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{R}^n$ *with* $A \subseteq \bigcup_{i \in \mathbb{N}} A_i$ *that* $v^n(A) \leqslant \sum_{i=1}^{\infty} v^n(A_i)$

Using the above with 18.33 the theorem is proved. □

We are now ready to define a σ-algebra and a measure on $\mathbb{R}^n$ using the Carathéodory theorem (see 18.34)

Definition 18.72. *Let* $\varphi^n(A) = \inf(\{\sum_{i=1}^{\infty} v^n(R_i) | \{R_i\}_{i \in \mathbb{N}} \subseteq \mathcal{R}^n \text{ where } A \subseteq \bigcup_{i \in \mathbb{N}} R_i\})$ *be the outer measure from the previous lemma then the* σ*-algebra defined by the Carathéodory method* $\{E \in \mathcal{P}(X) | \forall A \in \mathcal{P}(X) \text{ we have } \varphi^n(A) = \varphi^n(A \bigcap E) + \varphi^n(A \setminus E)\} = \{E \in \mathcal{P}(X) | \forall A \in \mathcal{P}(X) \text{ we have } \varphi^n(A) \geqslant \varphi^n(A \bigcap E) + \varphi^n(A \setminus E)\}$ *is called the set of Lebesgue measurable sets and is noted by* $\mathcal{L}^n$*. The measure* λ^n: $\mathcal{L}^n \rightarrow [0, \infty]$ *defined by* $\lambda^n = (\varphi^n)_{|\mathcal{L}^n}$ *is called the Lebesgue measure. This makes* $\langle \mathbb{R}^n, \mathcal{L}^n, \lambda^n \rangle$ *a measure space.*

Just as with the real numbers we proceed now to prove that the Borel algebra defined by the open sets in $\mathbb{R}^n$ is included in $\mathcal{L}^n$. First we prove that the set of half open spaces are Lebesgue measurable.

Lemma 18.73. *Let* $n \in \mathbb{N}$*,* $x \in \mathbb{R}$ *and* $i \in \{1, \ldots, n\}$ *then* $H_x^i = \{y \in \mathbb{R}^n | y_i < x\}$ *is Lebesgue measurable or* $H_\varepsilon^i \in \mathcal{L}^n$

Proof. First we prove that

$$\forall R \in \mathcal{R}^n \text{ we have } v^n(R) = v^n\big(R \bigcap H_x^i\big) + v^n(R \setminus H_x^i) \tag{18.126}$$

Proof. We must consider the following cases for R

$\boldsymbol{R \subseteq H_x^i}$**.** then

$$\begin{aligned} v^n\big(R \bigcap H_x^i\big) + v^n(R \setminus H_x^i) &= v^n(R) + v^n(\emptyset) \\ &= v^n(R) \end{aligned}$$

$\boldsymbol{R \bigcap H_x^i = \emptyset}$. then

$$\begin{aligned} v^n\left(R \bigcap H_x^i\right) + \lambda v^n(R \setminus H_x^i) &= v^n(\emptyset) + v^n(R) \\ &= v^n(R) \end{aligned}$$

$\boldsymbol{R \nsubseteq H_x^i \wedge R \bigcap H_x^i \neq \emptyset}$. then $R \neq \emptyset$ so there exists a $a, b \in \mathbb{R}^n$ with $a < b$ such that $R = [a, b[$. Consider now for x the following cases

$\boldsymbol{b_i \leqslant x}$. let $y \in R$ then $a_i \leqslant y_i < b_i \leqslant x \Rightarrow y_i < x$ so that $y \in H_x^i$, hence $R \subseteq H_x^i$ contradicting $R \nsubseteq H_x^i$

$\boldsymbol{x \leqslant a_i}$. as $R \bigcap H_x^i \neq \emptyset$ there exists a $y \in R \bigcap H_x^i$ so that $a_i \leqslant y_i < b_i$ and $y_i < x \leqslant a_i$ giving the contradiction $y_i < y_i$

So we must have that $a_i < x < b_i$, then we have

$$\begin{aligned} y \in R \bigcap H_x^i &\Leftrightarrow \forall j \in \{1, \ldots, n\} \setminus \{i\}\ a_j \leqslant y_j < b_j \wedge a_i \leqslant y_i < b_i \wedge y_i < x \\ &\underset{x < b_j}{\Leftrightarrow} \forall j \in \{1, \ldots, n\} \setminus \{i\}\ a_j \leqslant y_j < b_j \wedge a_i \leqslant y_i < x \\ &\Leftrightarrow y \in [a, \bar{b}[\text{ where } \bar{b}_j = \begin{cases} x \text{ if } j = i \\ b_i \text{ if } j \in \{1, \ldots, n\} \setminus \{i\} \end{cases} \end{aligned}$$

so that $R \bigcap H_x^i \in \mathcal{R}^n$ and $v^n(R \bigcap H_x^i)$ exists where

$$\begin{aligned} v^n\left(R \bigcap H_x^i\right) &= \prod_{j \in \{1,, \ldots, n\}} (\bar{b}_j - a_j) \\ &= \left(\prod_{j \in \{1, \ldots, n\} \setminus \{i\}} (b_j - a_j)\right) \cdot (x - a_i) \end{aligned} \tag{18.127}$$

Further we have

$$\begin{aligned} y \in R \setminus H_x^i &\Leftrightarrow \forall j \in \{1, \ldots, n\} \setminus \{i\}\ a_j \leqslant y_j < b_j \wedge a_i \leqslant y_i < b_i \wedge x \leqslant y_i \\ &\underset{a_i < x}{\Leftrightarrow} \forall j \in \{1, \ldots, n\} \setminus \{i\}\ a_j \leqslant y_j < b_j \wedge x \leqslant y_i < b_i \\ &\Leftrightarrow y \in [\bar{a}, b[\text{ where } \bar{a}_j = \begin{cases} x \text{ if } j = i \\ a_j \text{ if } j \in \{1, \ldots, n\} \setminus \{i\} \end{cases} \end{aligned}$$

so that $R \setminus H_x^i \in \mathcal{R}^n$ and $v^n(R \setminus H_x^i)$ exists where

$$\begin{aligned} v^n(R \setminus H_x^i) &= \prod_{j \in \{1, \ldots, n\}} (b_j - \bar{a}_j) \\ &= \left(\prod_{j \in \{1, \ldots, n\} \setminus \{i\}} (b_j - a_j)\right) \cdot (b_i - x) \end{aligned} \tag{18.128}$$

So we have

$$\begin{aligned}
v^n\left(R\bigcap H_x^i\right)+v^n(R\setminus H_x^i) &\underset{(18.127)\text{ and }(18.128)}{=} \left(\prod_{j\in\{1,\ldots,n\}\setminus\{i\}}(b_j-a_j)\right)\cdot(x-a_i)+\left(\prod_{j\in\{1,\ldots,n\}\setminus\{i\}}(b_j-a_j)\right)\cdot(b_i-x)\\
&= \left(\prod_{j\in\{1,\ldots,n\}\setminus\{i\}}(b_j-a_j)\right)\cdot[x-a_i+b_i-x]\\
&= \left(\prod_{j\in\{1,\ldots,n\}\setminus\{i\}}(b_j-a_j)\right)\cdot(b_i-a_i)\\
&= \prod_{j\in\{1,\ldots,n\}}(b_j-a_j)\\
&= v^n(R)
\end{aligned}$$

so in all cases we have $v^n(R\bigcap H_x^i)+v^n(R\setminus H_x^i)=v^n(R)$ proving (18.126) $\square$

Let now $A\in\mathcal{P}(\mathbb{R}^n)$ and take $\varepsilon>0$ then from the definition of φ^n and the infinum there exists a $\{R_j\}_{j\in\mathbb{N}}\subseteq\mathcal{R}^n$ with $A\subseteq\bigcup_{j\in\mathbb{N}}R_i$ such that

$$\varphi^n(A)\leqslant\sum_{j=1}^{\infty}v^n(R_j)<\varphi^n(A)+\varepsilon. \tag{18.129}$$

Hence $A\bigcap H_x^i\subseteq(\bigcup_{j\in\mathbb{N}}R_j)\bigcap H_x^i \underset{1.107}{=} \bigcup_{j\in\mathbb{N}}(R_j\bigcap H_x^i)$ and $A\setminus H_x^i\subseteq(\bigcup_{j\in\mathbb{N}}R_j)\setminus H_x^i{}_{1.108}\bigcup_{j\in\mathbb{N}}(R_j\setminus H_x^i)$. Hence using the definition of a outer measure we have

$$\begin{aligned}
\varphi^n\left(A\bigcap H_x^i\right)+\varphi^n(A\setminus H_x^i) &\leqslant \varphi^n\left(\bigcup_{j\in\mathbb{N}}(R_j\bigcap H_x^i)\right)+\varphi^n\left(\bigcup_{j\in\mathbb{N}}(R_j\setminus H_x^i)\right)\\
&\leqslant \sum_{j=1}^{\infty}\varphi^n\left(R_j\bigcap H_x^i\right)+\sum_{j=1}\varphi^n(R_j\setminus H_x^i)\\
&\underset{17.114}{=} \sum_{j=1}^{\infty}\left(\varphi^n\left(R_j\bigcap H_x^i\right)+\varphi^n(R_j\setminus H_x^i)\right)\\
&\underset{(18.126)}{=} \sum_{j=1}^{\infty}\varphi^n(R_j)\\
&\leqslant_{(18.129)} \varphi^n(A)+\varepsilon
\end{aligned}$$

which by the fact that ε is chosen arbitrary and 9.56 proves that

$$\varphi^n\left(A\bigcap H_x^i\right)+\varphi^n(A\setminus H_x^i)\leqslant\varphi^n(A)$$

and as A was chosen arbitrary that

$$H_x^i \in \mathcal{L}^n \qquad \square$$

Corollary 18.74. *Let $n \in \mathbb{N}$ then $\sigma[\{H_x^i | i \in \{1, \ldots, n\} \text{ and } x \in \mathbb{R}\}] \subseteq \mathcal{L}^n$*

Proof. Using 18.73 we have that $\{H_x^i | i \in \{1, \ldots, n\} \text{ and } x \in \mathbb{R}\} \subseteq \mathcal{L}^n$ then as $\mathcal{L}^n$ is a σ-algebra and $\sigma[\{H_x^i | i \in \{1, \ldots, n\} \text{ and } x \in \mathbb{R}\}]$ is the smallest σ-algebra covering $\{H_x^i | i \in \{1, \ldots, n\} \text{ and } x \in \mathbb{R}\}$ the corollary follows. $\square$

Next we extend the above to the half open intervals.

Lemma 18.75. *Let $n \in \mathbb{N}$ then $\sigma[\mathcal{R}^n] = \sigma[\{H_x^i | i \in \{1, \ldots, n\} \text{ and } x \in \mathbb{R}\}] \subseteq \mathcal{L}^n$*

Proof. First given $a \in \mathbb{R}^n$ take $]-\infty, a[\underset{\text{def}}{=} \{x \in \mathbb{R}^n | x < a\}$ then

$$\begin{aligned} x \in \{x \in \mathbb{R}^n | x < a\} &\Leftrightarrow \forall i \in \{1, \ldots, n\}\, x_i < a_i \\ &\Leftrightarrow \forall i \in \{1, \ldots, n\}\, x \in \{x \in \mathbb{R}^n | x_i < a_i\} \\ &\Leftrightarrow \forall i \in \{1, \ldots, n\}\, x \in H_{a_i}^i \end{aligned}$$

which proves that
we prove that

$$\forall a \in \mathbb{R}^n\,]-\infty, a[\underset{\text{def}}{=} \{x \in \mathbb{R}^n | x < a\} = \bigcap_{j \in \{1, \ldots, n\}} H_{x_j}^j \tag{18.130}$$

which by 18.8 proves that $]-\infty, a[= \bigcap_{j \in \{1, \ldots, n\}} H_{a_j}^j \in \sigma[\{H_x^i | i \in \{1, \ldots, n\} \text{ and } x \in \mathbb{R}\}]$ (using 18.8) which proves that

$$\forall a \in \mathbb{R}^n \quad]-\infty, a[\in \sigma[\{H_x^i | i \in \{1, \ldots, n\} \text{ and } x \in \mathbb{R}\}] \tag{18.131}$$

Next if $a, b \in \mathbb{R}^n$ then

$$\begin{aligned} x \in]-\infty, b[\setminus]-\infty, a[&\Leftrightarrow \forall i \in \{1, \ldots, n\}\, x_i < b_i \wedge \neg(x_i < a_i) \\ &\Leftrightarrow \forall i \in \{1, \ldots, n\}\, x_i < b_i \wedge a_i \leqslant x_i \\ &\Leftrightarrow a \leqslant x \wedge x < b \\ &\Leftrightarrow x \in [a, b[\end{aligned}$$

hence $[a, b[=]-\infty, b[\setminus]-\infty, a[\in \sigma[\{H_x^i | i \in \{1, \ldots, n\} \text{ and } x \in \mathbb{R}\}]$ (using 18.130 and 18.8). So we have that $\mathcal{R}^n \subseteq \sigma[\{H_x^i | i \in \{1, \ldots, n\} \text{ and } x \in \mathbb{R}\}]$. Hence as $\sigma[\mathcal{R}^n]$ is the smallest σ-algebra containing $\mathcal{R}^n$ we have

$$\sigma[\mathcal{R}^n] \subseteq \sigma[\{H_x^i | i \in \{1, \ldots, n\} \text{ and } x \in \mathbb{R}\}] \tag{18.132}$$

Next let $i \in \mathbb{N}$ and $x \in \mathbb{R}$ define then $\{[a_j, b_j[\}_{j \in \mathbb{N}} \subseteq \mathcal{R}^n$ where $\forall k \in \{1, \ldots, n\}$ we have $(a_j)_k = -j$ and $(b_j)_k = \left\{ \begin{smallmatrix} x \text{ if } k = i \\ j \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \end{smallmatrix} \right.$ then we have

$$\begin{aligned} y \in \bigcup_{j \in \mathbb{N}} [a_j, b_j[&\Rightarrow y_i \leqslant x \\ &\Rightarrow y \in H_x^i \end{aligned}$$

proving that

$$\bigcup_{j\in\mathbb{N}} [a_j, b_j[\subseteq H_x^i \tag{18.133}$$

Let $y\in H_x^i$ then for all $k\in\{1,\ldots,n\}$ there exists a $n_k\in\mathbb{N}$ such that $|y_k|<n_k$ (see 9.55) so that $y_k, -y_x<n_k \Rightarrow -n_k<y_k<n_k$. If we take $N=\max(\{n_k|k\in\{1,\ldots,n\}\})$ then $\forall k\in\{1,\ldots,n\}$ we have $(a_N)_k=-N<y_k$ and $\forall k\in\{1,\ldots,n\}\setminus\{i\}$ we have $y_k<N=(b_N)_k$ and $y_i<x=(b_N)_i$ proving that $y\in[a_N, b_N[$. Hence $H_x^i\subseteq\bigcup_{j\in\mathbb{N}}[a_j, b_j[$ which together with (18.133) proves that $H_x^i=\bigcup_{J\in\mathbb{N}}[a_j, b_j[\in\sigma[\mathcal{R}^n]$ (because of the definition of a σ-algebra). So we have proved that $\{H_x^i|i\in\{1,\ldots,n\}$ and $x\in\mathbb{R}\}\subseteq\sigma[\mathcal{R}^n]$ or

$$\sigma[\{H_x^i|i\in\{1,\ldots,n\} \text{ and } x\in\mathbb{R}\}]\subseteq\sigma[\mathcal{R}^n] \tag{18.134}$$

proving together with (18.132) that

$$\sigma[\mathcal{R}^n]=\sigma[\{H_x^i|i\in\{1,\ldots,n\} \text{ and } x\in\mathbb{R}\}]\subseteq_{18.74}\mathcal{L}^n \qquad \square$$

Now we introduce the concept of dyadic cubes to prove that every open sets is a countable union of half open intervals. Similar to what we have done in 18.55.

Definition 18.76. *Let $n, m\in\mathbb{N}$ and $z\in\mathbb{Z}^n$ then $\left[\frac{z}{2^n}, \frac{z+1}{2^n}\right[=\prod_{i\in\{1,\ldots,n\}}\left[\frac{z_i}{2^n}, \frac{z_i+1}{2^n}\right[$ is a dyadic cube of order m. The set of dyadic cubes of order m is noted as $\mathcal{D}_m^n$ so $\mathcal{D}_m^n=\left\{\left[\frac{z_i}{2^m}, \frac{z_i+1}{2^m}\right[|z\in\mathbb{Z}^n\right\}$. The set of all dyadic intervals is noted as $\mathcal{D}^n$ so $\mathcal{D}^n=\bigcup_{i\in\mathbb{N}}\mathcal{D}_i^n$.*

Note 18.77. As $\frac{z}{2^n}\in\left[\frac{z}{2^n}, \frac{z+1}{2^n}\right[$ it follows that dyadic cubes are always non empty.

Dyadic cubes have the same properties as Dyadic intervals (see 18.54)

Lemma 18.78. *Let $n\in\mathbb{N}$ then dyadic cubes have the following properties*

1. *$\forall m\in\mathbb{N}$ and $\forall R, Q\in\mathcal{D}_m^n$ with $R\neq Q$ we have $R\bigcap Q=\emptyset$*
2. *If $k, l\in\mathbb{N}$ with $k\leqslant l$ we have $\forall R\in\mathcal{D}_l^n$, $\forall Q\in\mathcal{D}_k^n$ we have either $R\subseteq Q$ or $Q\bigcap R=\emptyset$*
3. *$\forall m\in\mathbb{N}$ we have $\mathbb{R}^n=\bigsqcup_{R\in\mathcal{D}_m^n}R$ (a pairwise disjoint union)*
4. *$\forall m\in\mathbb{N}$ and $\forall R\in\mathcal{D}_m^n$ we have $v^n(R)=2^{-n\cdot m}$*

Proof.

1. Let $m\in\mathbb{N}$ and $R, Q\in\mathcal{D}_m^n$ with $R\neq Q$ then $R=\prod_{i\in\{1,\ldots,n\}}\left[\frac{z_i}{2^m}, \frac{z_i+1}{2^m}\right[$ and $Q=\prod_{i\in\{1,\ldots,n\}}\left[\frac{w_i}{2^m}, \frac{w_i+1}{2^m}\right[$ where $z, w\in\mathbb{Z}^n$. As $R\neq Q$ there exists a $i\in\{1,\ldots,n\}$ such that $\left[\frac{z_i}{2^m}, \frac{z_i+1}{2^m}\right[\neq\left[\frac{w_i}{2^m}, \frac{q_i+1}{2^m}\right[$, using then 18.54 (2) we have that $\left[\frac{z_i}{2^m}, \frac{z_i+1}{2^m}\right[\bigcap\left[\frac{w_i}{2^m}, \frac{q_i+1}{2^m}\right[=\emptyset$. So using 5.102 we have that

$$R\bigcap Q=\emptyset$$

2. Let $k,l\in\mathbb{N}$ with $k\leqslant l$ and $R\in\mathcal{D}_l^n$, $Q\in\mathcal{D}_k^n$ then there exists $z,w\in\mathbb{R}^n$ so that $R=\left[\frac{z}{2^l},\frac{z+1}{2^l}\right[$ and $Q=\left[\frac{w}{2^k},\frac{w+1}{2^k}\right[$. If $R\subseteq Q$ then as $R\neq\emptyset$ we have $R\bigcap Q=R\neq\emptyset$. If $R\nsubseteq Q$ we have either

   $k=l$. then as $R\nsubseteq Q$ we have that $R\neq Q$ and thus by (1) that $R\bigcap Q=\emptyset$

   $k<l$. then we have either

   $\boldsymbol{\exists i\in\{1,\ldots,n\}\vdash\left[\frac{z_i}{2^l},\frac{z_i+1}{2^l}\right[\bigcap\left[\frac{w_i}{2^k},\frac{w_i+1}{2^k}\right[=\emptyset}$. then using 5.102 it follows that $R\bigcap Q=\emptyset$

   $\boldsymbol{\forall i\in\{1,\ldots,n\}\vdash\left[\frac{z_i}{2^l},\frac{z_i+1}{2^l}\right[\bigcap\left[\frac{w_i}{2^k},\frac{w_i+1}{2^k}\right[\neq\emptyset}$. then using 18.54 (1) we have that $\forall i\in\{1,\ldots,n\}$ $\left[\frac{z_i}{2^l},\frac{z_i+1}{2^l}\right[\subseteq\left[\frac{w_i}{2^k},\frac{w_i+1}{2^k}\right[$ so that using 5.102 it follows that $R\subseteq Q$ contradiction $R\nsubseteq Q$. So this case can not occur.

   so we must have $R\bigcap Q=\emptyset$.

3. Let $m\in\mathbb{N}$. If $x\in\mathbb{R}^n$ then $\forall i\in\{1,\ldots,n\}$ we have $x_i\in\mathbb{R}\underset{18.54\,(3)}{\Rightarrow}\exists z_i\in\mathbb{Z}$ such that $x_i\in\left[\frac{z_i}{2^m},\frac{z_i+1}{2^m}\right[$. Hence if we take $z=\{z_1,\ldots,z_n\}\in\mathbb{Z}^n$ we have that $x\in\left[\frac{z}{2^m},\frac{z+1}{2^m}\right[=\prod_{i\in\{1,\ldots,n\}}\left[\frac{z_i}{2^m},\frac{z_i+1}{2^m}\right[\in\mathcal{D}_m^n$ proving that $\mathbb{R}^n\subseteq\bigcup_{R\in\mathcal{D}_m^n}R$. As also $\forall R\in\mathcal{D}_m^n$ we have $R\subseteq\mathbb{R}^n$ it follows that $\bigcup_{R\in\boldsymbol{D}_m^n}R\subseteq\mathbb{R}^n$. So also using (1) we have
$$\mathbb{R}^n=\bigsqcup_{R\in\mathcal{R}_m^n}R$$

4. If $R\in\mathcal{R}_m^n$ then $\exists z\in\mathbb{Z}^n$ such that $R=\prod_{i\in\{1,\ldots,n\}}\left[\frac{z_i}{2^m},\frac{z_i+1}{2^m}\right[$ so that
$$v^n(R)=\prod_{i\in\{1,\ldots,n\}}\left(\frac{z_i+1-z_i}{2^m}\right)=\prod_{i\in\{1,\ldots,n\}}\frac{1}{2^m}=\frac{1}{2^{n\cdot m}}$$ □

Next we prove that the set of Dyadic cubes is denumerable

Lemma 18.79. *Let $n,m\in\mathbb{N}$ then $\mathcal{D}_m^n$ is denumerable and in addition $\mathcal{D}^n$ is also denumerable*

Proof. Let $n,m\in\mathbb{N}_0$. As $\mathbb{Z}$ is denumerable (see 8.54) we have by 5.101 that $\mathbb{Z}^n$ is also denumerable. Define now $\beta:\mathbb{Z}^n\to\mathcal{D}_m^n$ by $\beta(z)=\left[\frac{z}{2^m},\frac{z+1}{2^m}\right[$ then β is a bijection

injectivity. If $\beta(z)=\beta(w)$ then $\left[\frac{z}{2^m},\frac{z+1}{2^m}\right[=\left[\frac{w}{2^m},\frac{w+1}{2^m}\right[\underset{\text{dyadic intervals ar not empty and 18.66}}{\Rightarrow}\frac{z}{2^m}=\frac{w}{2^m}$ proving that $z=w$

surjectivity. If $D\in\mathcal{D}_m^n$ then $\exists z\in\mathbb{Z}^n$ such that $D=\left[\frac{z}{2^m},\frac{z+1}{2^m}\right[=\beta(z)$

so $\mathcal{D}_m^n$ is denumerable. Further using 5.60 we have that $\mathcal{D}^n=\bigcup_{m\in\mathbb{N}}\mathcal{D}_m^n$ is also denumerable. □

We decompose now open sets in $\mathbb{R}^n$ as a countable union of dyadic cubes.

Lemma 18.80. *Let $n \in \mathbb{N}$ and $\mathcal{T}_{\mathbb{R}^n}$ the canonical topology defined on $\mathbb{R}^n$ (see 12.80) then if $\emptyset \neq U \in \mathcal{T}_{\mathbb{R}^n}$ is a non empty open set then there exists a denumerable pairwise disjoint family $\{D_i\}_{i \in \mathbb{N}} \subseteq \mathcal{D}^n$ such that $U = \bigsqcup_{i \in \mathbb{N}} D_i$*

Proof. Let $\emptyset \neq U \in \mathcal{T}_{\mathbb{R}^n}$ be a non empty open set in $\mathbb{R}^n$. We use the following definitions for $m \in \mathbb{N}$

$$\mathcal{S}_m = \{D \in \mathcal{D}^n_m | D \subseteq U\} \subseteq \mathcal{D}^n_m \tag{18.135}$$

$$\mathcal{T}_m = \left\{ \begin{array}{l} \mathcal{S}_0 \text{ if } m = 0 \\ \{D \in \mathcal{S}_n | \forall i \in \{1, \ldots, n-1 \text{ we have } \forall P \in \mathcal{T}_i \text{ that } P \bigcap D = \emptyset\}\} \text{ if } m > 0 \end{array} \right. \tag{18.136}$$

from the above definitions it quickly follows that

$$\mathcal{T}_m \subseteq \mathcal{S}_m \subseteq \mathcal{D}^n_m \subseteq \mathcal{D}^n$$

finally we define

$$\mathcal{T} = \bigcup_{m \in \mathbb{N}} \mathcal{T}_m \subseteq \mathcal{D}^n \tag{18.137}$$

First as $\forall D \in \mathcal{T}$ we have that $\exists i \in \mathbb{N}$ such that $D \in \mathcal{T}_i \subseteq S_i \Rightarrow D \subseteq U$ it follows that

$$\bigcup_{D \in \mathcal{T}} D \subseteq U \tag{18.138}$$

For the opposite inclusion, take $x \in U$ then as U is open there exist a $B_{\|\|}(x, \varepsilon)$ such that $x \in B_{\|\|}(x, \varepsilon) \subseteq U$. Using 9.55 there exists a $n_x \in \mathbb{N}$ such that $0 < \frac{1}{n_x} < \varepsilon$. Using 18.78 (3) there exists a $D = \prod_{i \in \{1, \ldots, n\}} \left[\frac{z_i}{2^{n_x}}, \frac{z_i + 1}{2^{n_x}}\right[\in \mathcal{D}^n_m$ such that $x \in D$. So $\forall i \in \{1, \ldots, n\}$ we have $\frac{z_i}{2^{n_x}} \leqslant x_i < \frac{z_i + 1}{2^{n_x}}$. If $y \in D$ then $\forall i \in \{1, \ldots, n\}$ we have $\frac{z_i}{2^{n_x}} \leqslant y_i < \frac{z_i + 1}{2^{n_x}}$, from this it follows that $\frac{z_i}{2^{n_x}} - \frac{z_i + 1}{2^{n_x}} \leqslant x_i - y_i < \frac{z_i + 1}{2^{n_x}} - \frac{z_i}{2^{n_x}} \Rightarrow -\frac{1}{2^{n_x}} \leqslant x_i - y_i < \frac{1}{2^{n_x}} \Rightarrow -\varepsilon < x_i - y_i < \varepsilon \Rightarrow |x_i - y_i| < \varepsilon \Rightarrow \max(\{|x_i - y_i|| i \in \{1, \ldots, n\}\}) < \varepsilon$ proving that $y \in B_{\|\|}(x, \varepsilon)$. So we have $x \in D \subseteq B_{\|\|}(x, \varepsilon) \subseteq U$ or $D \in \mathcal{S}_{n_x}$. If we define then $\mathcal{K}_x = \{i \in \{1, \ldots, n_x\} | \exists D \in \mathcal{S}_i \text{ such that } x \in D\}$ then $n_x \in \mathcal{K}_x$ so that $\mathcal{K}_x \neq \emptyset$, hence $m_x = \min(\mathcal{K}_x)$. We have now two cases to consider

- $\boldsymbol{m_x = 1.}$ then $\exists D \in \mathcal{S}_1 = \mathcal{T}_1 \subseteq \mathcal{T}$ such that $x \in D \subseteq \bigcup_{D \in \mathcal{T}} D$
- $\boldsymbol{1 < m_x.}$ then $\exists D \in \mathcal{S}_{m_x}$(such that $x \in D$) and for $i \in \{1, \ldots, m_x - 1\}$ we have $\forall R \in \mathcal{S}_i$ that $x \notin R$. Now as $i < m_x$ we have by 18.78 (2) that either $D \subseteq R$, but then as $x \in D \subseteq R$ we contradict $x \notin R$, or $D \bigcap R = \emptyset$. So $\forall i \in \{1, \ldots, m_x - 1\}$ we have $\forall R \in \mathcal{S}_i$ we have $R \bigcap D = \emptyset$ proving that $D \in \mathcal{T}_{m_x} \subseteq \mathcal{T}$ (see (18.136)) hence $x \in D \subseteq \bigcup_{D \in \mathcal{T}} D$.

As we have chosen $x \in U$ arbitrary it follows that $U \subseteq \bigcup_{D \in \mathcal{T}} D$ which together with (18.138) proves

$$U = \bigcup_{D \in \mathcal{T}} D \tag{18.139}$$

Let $D, E \in \mathcal{T}$ with $D \neq E$ then there exists $k, l \in \mathbb{N}$ such that $D \in \mathcal{T}_k \subseteq \mathcal{D}^n_k$, $E \in \mathcal{T}_l \subseteq \mathcal{D}^n_l$. For k, l we have the following possibilities

- $\boldsymbol{k = l.}$ then using 18.78 we have $D \bigcap E = \emptyset$

$\boldsymbol{k<l}$**.** then using (18.136) we have $D\bigcap E=\emptyset$

$\boldsymbol{l<k}$**.** then using (18.136) we have $D\bigcap E=\emptyset$

combining this with (18.139) proves that

$$U=\bigsqcup_{D\in\mathcal{T}} D \tag{18.140}$$

Assume now that $\mathcal{T}$ is finite. Then as $U\neq\emptyset$ we must have that $\mathcal{T}\neq\emptyset$ then $m=\min\{\text{begin}(D)_1|D\in\mathcal{T}\}$ is well defined and thus there exists a $R=[a,b[\in\mathcal{T}$, such that $a_1=m$. As $a\in R\subseteq U$ a open set there exists a $\varepsilon>0$ such that $a\in B_{\|\|}(a,\varepsilon)\subseteq U$. As $a-\frac{\varepsilon}{2}\subseteq B_{\|\|}(a,\varepsilon)\subseteq U$ there exists a $D=[a',b'[\in\mathcal{T}$ such that $a-\frac{\varepsilon}{2}\in D$ then $m\leqslant a_1'=\text{begin}(D)_1\leqslant a_1-\frac{\varepsilon}{2}<a_1=m$ a contradiction. Se we must have that $\mathcal{T}$ is infinite. As $\mathcal{T}\subseteq\mathcal{D}^n$ a denumerable set(see 18.79) it follows from 5.55 that

$$\mathcal{T}\text{ is denumerable} \tag{18.141}$$

So there exist a bijection $\beta:\mathbb{R}\rightarrow\mathcal{T}$, define then $\{D_i\}_{i\in\mathbb{N}}$ by $D_i=\beta(i)$ then we have using 2.64

$$U=\bigsqcup_{i\in\mathbb{N}} D_i$$

proving the lemma. □

Actually if we don't need to write a open set as union of **disjoint** non empty intervals we do not need to use dyadic intervals. This is illustrated in the following lemma.

Lemma 18.81. *Let $n\in\mathbb{N}$ and define $\mathcal{I}^n=\{]a,b]|a,b\in\mathbb{R}^n\,\text{with}\,a\leqslant b\}$ then $\forall U\in\mathcal{T}_{\mathbb{R}^n}$ there exists a $\{R_i\}_{i\in\mathbb{N}}\subseteq\mathcal{I}^n$ such that $\emptyset\neq U=\bigcup_{i\in\mathbb{N}}R_i$*

Proof. Take $U\in\mathcal{T}$ and define

$$\mathcal{U}=\{(q,q')\in\mathbb{Q}^n\times\mathbb{Q}^n|q\leqslant q'\wedge]q,q']\subseteq U\ \}\subseteq\mathbb{Q}^n\times\mathbb{Q}^n \tag{18.142}$$

Using 9.38, 5.69 and 5.66 we have that

$$\mathcal{U}\text{ is countable} \tag{18.143}$$

By the definition of $\mathcal{U}$ it is clear that

$$\bigcup_{(q,q')\in\mathcal{U}}]q,q']\subseteq U \tag{18.144}$$

Let $x\in U$ a open set there exists a $\delta_x>0$ such that $x\in B_{\|\|}(x,\delta_x)=\{y\in\mathbb{R}^n|\max(\{|x_i-y_i||i\in\{1,\ldots,n\}\})<\delta_x\}$. Given $i\in\{1,\ldots,n\}$ there exists by 9.57 a q_i, $q_i'\in\mathbb{Q}$ such that $x_i-\delta_x<q_i<x_i<q_i'<x_i+\delta_x$. Define then $q=(q_1,\ldots,q_n)$, $q'=(q_1',\ldots,q_n')\in\mathbb{Q}^n$ then $x\in]q,q']$ and if $y\in]q,q']$ we have $\forall i\in\{1,\ldots,n\}$ that $x_i-\delta_x<q<y\leqslant q'<x_i+\delta_x\Rightarrow x_i-y_i<\delta_x, y_i-x_i<\delta_x\Rightarrow|x_i-y_i|<\delta_x$ so that $\|x-y\|=\max(\{|x_i-y_i||i\in\{1,\ldots,n\}\})<\delta_x$ proving that $y\in B_{\|\|}(x,\delta_x)$. So $x\in]q,q']\subseteq U$ hence $(q,q')\in\mathcal{U}$ or $x\in\bigcup_{(q,q')}]q,q']$ proving that $U\subseteq\bigcup_{(q,q')}]q,q']$ and thus by (18.144) we have

$$U=\bigcup_{(q,q')\in\mathcal{U}}]q,q'] \tag{18.145}$$

Now as $\mathcal{U}$ is countable we have either

$\mathcal{U}$ is finite. Then as $\emptyset \neq U$ we have that $\mathcal{U} \neq \emptyset$ and thus there exists a $m \in \mathbb{N}$ and a bijection $\beta: \{1, \ldots, m\} \rightarrow \mathcal{U}$. Define now $\{R_i\}_{i \in \mathbb{N}}$ by $R_i = \left\{ \begin{smallmatrix}]\beta(i)_1, \beta(i)_2] \text{ if } i \in \{1, \ldots, m\} \\ \emptyset =]1, 1] \end{smallmatrix} \right.$ then $\{R_i\}_{i \in \mathbb{N}} \subseteq \mathcal{I}^n$ and $\bigcup_{i \in \mathbb{N}} R_i \underset{1.109}{=} \bigcup_{i \in \{1, \ldots, m\}} R_i \underset{2.64}{=} \bigcup_{(q, q') \in \mathcal{U}}]q, q'] = U$

$\mathcal{U}$ is infinitely countable. Then there exists a bijection $\beta: \mathbb{N} \rightarrow \mathcal{U}$ and thus if we define $\{R_i\}_{i \in \mathbb{N}}$ by $R_i =]q_{\beta(i)}, q_{\beta(i)_2}]$ we have that $\bigcup_{i \in \mathbb{N}} R_i \underset{2.64}{=} \bigcup_{(q, q') \in \mathcal{U}}]q, q'] = U$

So in all case we have

$$\exists \{R_i\}_{i \in \mathbb{N}} \subseteq \mathcal{I}^n \text{ such that } \bigcup_{i \in \mathbb{N}} R_i = U$$

□

The next lemma allows us the write a half-open interval in $\mathbb{R}^n$ as a countable intersection of open sets.

Lemma 18.82. *Let $n \in \mathbb{N}$ then $\forall R \in \mathcal{R}^n$ with $\emptyset \neq R$ there exists a sequence $\{U_i\}_{i \in \mathbb{N}} \subseteq \mathcal{T}_{\mathbb{R}^n}$ of open sets such that $R = \bigcap_{i \in \mathbb{N}} U_i$*

Proof. As $R \neq \emptyset$ there exists $a, b \in \mathbb{R}$ such that $R = [a, b[= \prod_{j \in \{1, \ldots, n\}} [a_j, b_j[$. Define then $\{U_i\}_{i \in \mathbb{N}}$ by $U_i = \prod_{j \in \{1, \ldots, n\}} \left] a_j - \frac{1}{i}, b_j \right[$, then as $\left] a_j - \frac{1}{i}, b_j \right[\in \mathcal{T}_{\mathbb{R}}$ and $\mathcal{T}_{\mathbb{R}^n}$ is the product topology based on $\mathcal{T}_{\mathbb{R}}$ (see 12.80) it follows that $\{U_i\}_{i \in \mathbb{N}} \subseteq \mathcal{T}_{\mathbb{R}^n}$. Let $i \in \mathbb{N}$ then $\forall j \in \{1, \ldots, n\}$ we have $]a_j, b_j[\subseteq \left] a_j - \frac{1}{i}, b_j \right[$ so that using 5.102 we have $R = \prod_{j \in \{1, \ldots, n\}} [a_j, b_j[\subseteq \prod_{j \in \{1, \ldots, n\}} \left] a_j - \frac{1}{i}, b_j \right[= U_i$. This proves

$$R \subseteq \bigcap_{i \in \mathbb{N}} U_i \tag{18.146}$$

For the opposite inequality let $x \in \bigcap_{i \in \mathbb{N}} U_i$. Take $j \in \{1, \ldots, n\}$ then $\forall i \in \mathbb{N}$ we have $a_j - \frac{1}{i} < x_j < b_j$. Assume now that $x_j < a_j \Rightarrow 0 < a_j - x_j$ then using 9.55 there exists a $i \in \mathbb{N}$ such that $0 < \frac{1}{i} < a_j - x_j \Rightarrow x_j < a_j - \frac{1}{i}$ contradicting $a_j - \frac{1}{i} < x_j$ $\forall i \in \mathbb{N}$. Hence $\forall j \in \{1, \ldots, n\}$ we have $a_j \leqslant x_j < b_j$ which proves that $x \in \prod_{j \in \{1, \ldots, n\}} [a_j, b_j[$ or $\bigcap_{i \in \mathbb{N}} U_i \subseteq R$. Combining this with (18.146) gives finally

$$R = \bigcap_{i \in \mathbb{N}} U_i$$

□

We are now ready to prove that $\mathcal{B}[\mathbb{R}^n] = \sigma[\mathcal{R}^n]$

Lemma 18.83. $\mathcal{B}[\mathbb{R}^n] = \sigma[\mathcal{R}^n]$ *(where $\mathcal{B}[\mathbb{R}^n] = \mathcal{B}[\mathbb{R}^n, \mathcal{T}_{\mathbb{R}^n}] = \sigma[\mathcal{T}_{\mathbb{R}^n}]$ see 18.13)*

Proof. Let $U \in \mathcal{T}_{\mathbb{R}}$ then we have either

$U = \emptyset$. then $U \in \sigma[\mathcal{R}^n]$

$U \neq \emptyset$. then using 18.80 there exists a $\{D_i\}_{i \in \mathbb{N}} \subseteq \mathcal{R}^n$ such that $U = \bigcup_{i \in \mathbb{N}} D_i \in \sigma[\mathcal{R}^n]$ [as $\sigma[\mathcal{R}^n]$ is a σ-algebra] hence $U \in \sigma[\mathcal{R}^n]$

so $\mathcal{T}_{\mathbb{R}^n} \subseteq \sigma[\mathcal{R}^n]$ and thus

$$\sigma[\mathcal{T}_{\mathbb{R}^n}] \subseteq \sigma[\mathcal{R}^n] \tag{18.147}$$

Further if $R \in \mathcal{R}^n$ then we have either

$R = \emptyset$. then $R \in \sigma[\mathcal{T}_{\mathbb{R}^n}]$

$R \neq \emptyset$. then using 18.82 there exists a $\{U_i\}_{i \in \mathbb{N}} \subseteq \mathcal{T}_{\mathbb{R}^n}$ such that $R = \bigcap_{i \in \mathbb{N}} R_i \in \sigma[\mathcal{T}_{\mathbb{R}^n}]$ [because $\sigma[\mathcal{T}_{\mathbb{R}^n}]$ is a σ-algebra together with 18.8]

proving $\mathcal{R}^n \in \sigma[\mathcal{T}_{\mathbb{R}^n}]$ and thus $\sigma[\mathcal{R}^n] \subseteq \sigma[\mathcal{T}_{\mathbb{R}^n}]$. Combining this with (18.147) gives

$$\sigma[\mathcal{T}_{\mathbb{R}^n}] = \sigma[\mathcal{R}^n]$$

proving the lemma. □

A alternate formulation of $\mathcal{B}[\mathbb{R}^n]$ is expressed in the following lemma

Definition 18.84. *Let $n \in \mathbb{N}$ then given $i \in \{1, \ldots, n\}$ and $a \in \mathbb{R}$ we define the lower half space $]-\infty, a]_i$ by $]-\infty, a]_i = \{x \in \mathbb{R}^n | x_i \leqslant a\}$ and the set of all lower half spaces is noted by $\mathcal{H}^n$ so $\mathcal{H}^n = \{]-\infty, a]_i | a \in \mathbb{R}, i \in \{1, \ldots, n\}\}$*

Lemma 18.85. *Let $n \in \mathbb{N}$ then $\forall H \in \mathcal{H}^n$ we have that H is closed in $\mathcal{T}_{\mathbb{R}^n}$*

Proof. Let $H \in \mathcal{H}^n$ then there exists a $i \in \{1, \ldots, n\}$ and a $a \in \mathbb{R}$ such that $H = \{x \in \mathbb{R}^n | x_i \leqslant a\}$. Take now $x \in \mathbb{R}^n \backslash H$ then $a < x_i$ so that $\varepsilon = x_i - a_i > 0$ then $x \in B_{\|\|}(x, \varepsilon)$ and if $y \in B_{\|\|}(x, \varepsilon)$ we have that $x_i - y_i | x_i - y_i| \leqslant \max(\{|x_j - y_j| | j \in \{1, \ldots, n\}\}) < \varepsilon = x_i - a_i$ so that $-y_i < -a_i$ or $a_i < x_i$ hence $y \in H$ or $y \in \mathbb{R}^n \backslash H$. So for every $x \in \mathbb{R}^n \backslash H$ we found a open set $B_{\|\|(x,\varepsilon)}$ containing x which proves that $\mathbb{R}^n \backslash H$ is open and that H is closed. □

Lemma 18.86. *Let $n \in \mathbb{N}$ then we have that $\mathcal{B}[\mathbb{R}^n] = \sigma[\mathcal{H}^n]$*

Proof. As $\forall H \in \mathcal{H}^n$ we have that H is closed we have by 18.14 that $H \in \mathcal{B}[\mathbb{R}^n]$ so that $\mathcal{H}^n \subseteq \mathcal{B}[\mathbb{R}^n]$. and thus

$$\sigma[\mathcal{H}^n] \subseteq \mathcal{B}[\mathbb{R}^n] \tag{18.148}$$

Take $R \in \mathcal{I}^n$ (see 18.81) then there exists $a, b \in \mathbb{R}^n$ with $a \leqslant b$ such that $R =]a, b]$ then we have

$$\begin{aligned} x \in \bigcap_{i \in \{1, \ldots, n\}} (]-\infty, b_i[_i \backslash]-\infty, a_i]_i) &\Leftrightarrow \forall i \in \{1, \ldots, n\}\, x \in]-\infty, b_i]_i \backslash]-\infty, a_i]_i \\ &\Leftrightarrow \forall i \in \{1, \ldots, n\}\, x_i \leqslant b_i \wedge \neg(x_i \leqslant a_i) \\ &\Leftrightarrow \forall i \in \{1, \ldots, n\}\, x_i \leqslant b_i \wedge a_i < x_i \\ &\Leftrightarrow \forall i \in \{1, \ldots, n\}\, x_i \in]a_i, b_i] \\ &\Leftrightarrow x \in]a, b] \end{aligned}$$

proving that $]a, b] = \bigcap_{i \in \{1, \ldots, n\}} (]-\infty, b_i[_i \backslash]-\infty, a_i]_i)$. Using 18.8 we have that $\bigcap_{i \in \{1, \ldots, n\}} (]-\infty, b_i[_i \backslash]-\infty, a_i]_i) \in \sigma[\mathcal{H}^n]$ so that

$$\mathcal{I}^n \subseteq \sigma[\mathcal{H}^n] \tag{18.149}$$

Let now $U \in \mathcal{T}_{\mathbb{R}^n}$ then using 18.81 there exists a $\{R_i\}_{i\in\mathbb{N}} \subseteq \mathcal{I}^n \subseteq_{(18.149)} \sigma[\mathcal{H}^n]$ such that $U = \bigcup_{i\in\mathbb{N}} R_i$ which proves that $U \in \sigma[\mathcal{H}^n]$. So $\mathcal{T}_{\mathbb{R}^n} \subseteq \sigma[\mathcal{H}^n]$ hence $\sigma[\mathcal{T}_{\mathbb{R}^n}] \subseteq \sigma[\mathcal{H}^n]$ which together with (18.149) proves that

$$\mathcal{B}[\mathbb{R}^n] = \sigma[\mathcal{T}_{\mathbb{R}^n}] = \sigma[\mathcal{H}^n]$$ □

To summarize the above lemmas we have by 18.75,18.83 and 18.86 that

Theorem 18.87. *Let $n \in \mathbb{N}$ then we have that for $\mathcal{L}^n$ the set of Lebesgue measurable sets of $\mathbb{R}^n$ (see 18.72) we have*

$$\begin{aligned} \sigma[\{[a,b[\,|a,b \in \mathbb{R}^n \text{ with } a \leqslant b\}] &= \sigma[\mathcal{H}^n] \\ &= \mathcal{B}[\mathbb{R}^n] \\ &\subseteq \mathcal{L}^n \end{aligned}$$

constructing the measure space $\langle \mathbb{R}^n, \mathcal{B}[\mathbb{R}^n], \lambda^n_{|\mathcal{B}[\mathbb{R}^n]}\rangle$.

18.2 Integration

18.2.1 Measurable functions

As not all functions that we want to integrate over are not always defined on the whole measure space we will use partial functions (see 2.1) in this text. So if we say that $f: X \to Y$ is a partial function from X to Y then $f(x)$ is not defined for every $x \in X$ but only for $x \in \operatorname{dom}(f)$ where $\operatorname{dom}(f) \subseteq X$. We will also extensively use conegligible subsets, for this we use the following definitions.

Let's look at some partial functions defined based of other partial functions

Definition 18.88. *Let X be a set*

1. *Let $f: X \to \mathbb{R}$ a partial function and $C \subseteq X$ then $f_{|C}: X \to \mathbb{R}$ is the partial function with $\operatorname{dom}(f_{|C}) = \operatorname{dom}(f) \bigcap C$ and $\forall x \in \operatorname{dom}(f_{|C})$ we have $(f_{|C})(x) = f(x)$ [see 2.27]*
2. *Let $f: X \to \mathbb{R}$ be a partial function then $|f|: X \to \mathbb{R}$ is the partial function with $\operatorname{dom}(|f|) = \operatorname{dom}(f)$ and $\forall x \in \operatorname{dom}(|f|)$ $|f|(x) = |f(x)|$*
3. *Let $f: X \to \mathbb{R}$ a partial function, $c \in \mathbb{R}$ then $c \cdot f: X \to \mathbb{R}$ defined by $(c \cdot f)(x) = c \cdot f(x)$ $\forall x \in \operatorname{dom}(c \cdot f) = \operatorname{dom}(f)$.*
4. *Let $f: X \to \mathbb{R}$ a partial function then $|f|: X \to \mathbb{R}$ defined by $|f|(x) = |f(x)|$ $\forall x \in \operatorname{dom}(|f|) = \operatorname{dom}(f)$.*
5. *Let $f: X \to \mathbb{R}$, $g: X \to \mathbb{R}$ be partial functions then $f + g: X \to \mathbb{R}$ is the partial function defined by $(f+g)(x) = f(x) + g(x)$ $\forall x \in \operatorname{dom}(f+g) = \operatorname{dom}(f) \bigcap \operatorname{dom}(g)$*
6. *Let $f: X \to \mathbb{R}$, $g: X \to \mathbb{R}$ be partial functions then $f \cdot g: X \to \mathbb{R}$ is the partial function defined by $(f \cdot g)(x) = f(x) \cdot g(x)$ $\forall x \in \operatorname{dom}(f \cdot g) = \operatorname{dom}(f) \bigcap \operatorname{dom}(g)$.*

7. *Let* $f: X \rightarrow \mathbb{R}$ *a partial functions then* $\frac{1}{f}: X \rightarrow \mathbb{R}$ *is the partial function defined by* $\frac{1}{f}(x) = \frac{1}{f(x)}$ $\forall x \in \mathrm{dom}\left(\frac{1}{f}\right) = \{x \in \mathrm{dom}(f) | f(x) \neq 0\}$.

8. *Let* $f: X \rightarrow \mathbb{R}$, $g: X \rightarrow \mathbb{R}$ *be partial functions then* $\frac{f}{g}: X \rightarrow \mathbb{R}$ *is the partial function defined by* $\frac{f}{g}(x) = \frac{f(x)}{g(x)}$ $\forall x \in \mathrm{dom}(f+g) = \mathrm{dom}(f) \bigcap \{x \in \mathrm{dom}(g) | g(x) \neq 0\}$.

9. *Let* $n \in \mathbb{N}$ *and* $\{f_n: X \rightarrow \mathbb{R}\}_{n \in \mathbb{N}}$ *be partial functions then* $\sum_{i=1}^{n} f_i: X \rightarrow \mathbb{R}$ *is the partial function defined by* $(\sum_{i=1}^{n} f_i)(x) = \sum_{i=1}^{n} f_i(x)$ $\forall x \in \mathrm{dom}(\sum_{i=1}^{n} f_i) = \bigcap_{i \in \{1,\ldots,n\}} \mathrm{dom}(f_i)$.

10. *Let* $\{f_n: X \rightarrow \mathbb{R}\}_{n \in \mathbb{N}}$ *be partial functions then* $\lim_{n \rightarrow \infty} f_n : X \rightarrow \mathbb{R}$ *is the partial function defined by* $\mathrm{dom}\left(\lim_{n \rightarrow \infty} f_n\right) = \left\{x \in \bigcup_{n \in \mathbb{N}} \left(\bigcap_{i \in \{n,\ldots,\infty\}} \mathrm{dom}(f_i)\right) | \exists k \in \mathbb{N} \text{ such that } x \in \bigcap_{i \in \{k,\ldots,\infty\}} \mathrm{dom}(f_i) \wedge \{f_n(x)\}_{n \in \{k,\ldots\infty\}} \text{ has a limit in } \mathbb{R}\right\}$ *and* $\left(\lim_{n \rightarrow \infty} f_n\right)(x) = \lim_{n \rightarrow \infty} f_n(x)$ *where* $\lim_{n \rightarrow \infty} f_n(x)$ *is the limit of* $\{f_n(x)\}_{n \in \{k,\ldots,\infty\}}$ *where* $k \in \mathbb{N}$ *such that* $x \in \bigcap_{i \in \{k,\ldots,\infty\}} \mathrm{dom}(f_i)$ *and* $\{f_n(x)\}_{n \in \{k,\ldots,\infty\}}$ *has a limit.*

    Proof. *Of course we must prove that this function is well defined. So assume that there exists* k, l *such that* $x \in \bigcap_{i \in \{k,\ldots,\infty\}} \mathrm{dom}(f_i), \bigcap_{i \in \{l,\ldots,\infty\}} \mathrm{dom}(f_i)$ *and* $\{f_i(x)\}_{i \in \{k,\ldots,\infty\}}, \{f_i(x)\}_{i \in \{k,\ldots\infty\}}$ *has a limit. We may then assume that* $k \leqslant l$ *[otherwise exchange* k *an* l*]. Then* $x \in \bigcap_{i \in \{k,\ldots,\infty\}} \mathrm{dom}(f_i)$ *and using 12.304 we have that* $\{f_i(x)\}_{i \in \{k,\ldots,\infty\}}$ *and* $\{f_i(x)\}_{i \in \{l,\ldots,\infty\}}$ *have the same limit.* □

11. *Let* $\{f_n: X \rightarrow \mathbb{R}\}_{n \in \mathbb{N}}$ *be partial functions then* $\sup_{n \in \mathbb{N}} f_n$ *is the partial function defined by* $\left(\sup_{n \in \mathbb{N}} f_n\right)(x) = \sup\left(\{f_n(x) | n \in \mathbb{N}\}\right)$ $\forall x \in \mathrm{dom}\left(\sup_{n \in \mathbb{N}} f_n\right) = \{x \in \bigcap_{n \in \mathbb{N}} \mathrm{dom}(f_n) | \sup\left(\{f_n(x) | n \in \mathbb{N}\}\right) \in \mathbb{R}\}$.

12. *Let* $\{f_n: X \rightarrow \mathbb{R}\}_{n \in \mathbb{N}}$ *be partial functions then* $\inf_{n \in \mathbb{N}} f_n$ *is the partial function defined by* $\left(\inf_{n \in \mathbb{N}} f\right)(x) = \inf\left(\{f_n(x) | n \in \mathbb{N}\}\right)$ $\forall x \in \mathrm{dom}\left(\inf_{n \in \mathbb{N}} f\right) = \{x \in \bigcap_{n \in \mathbb{N}} \mathrm{dom}(f_n) | \inf\left(\{f_n(x) | n \in \mathbb{N}\}\right) \in \mathbb{R}\}$.

13. *Let* $\{f_n: X \rightarrow \mathbb{R}\}_{n \in \mathbb{N}}$ *be partial functions then* $\liminf_{n \rightarrow \infty} f_n$ *is the partial function defined by* $\left(\liminf_{n \rightarrow \infty} f_n\right)(x) = \liminf_{n \rightarrow \infty} f_n(x)$ $\forall x \in \mathrm{dom}(\sup\left(\{f_n | n \in \mathbb{N}\}\right)) = \left\{x \in \bigcap_{n \in \mathbb{N}} \mathrm{dom}(f_n) | \liminf_{n \rightarrow \infty} f_n(x) \in \mathbb{R}\right\}$.

14. *Let* $\{f_n: X \rightarrow \mathbb{R}\}_{n \in \mathbb{N}}$ *be partial functions then* $\limsup_{n \rightarrow \infty} f_n$ *is the partial function defined by* $\left(\limsup_{n \rightarrow \infty} f_n\right)(x) = \limsup_{n \rightarrow \infty} f_n(x)$ $\forall x \in \mathrm{dom}(\sup\left(\{f_n | n \in \mathbb{N}\}\right)) = \left\{x \in \bigcap_{n \in \mathbb{N}} \mathrm{dom}(f_n) | \limsup_{n \rightarrow \infty} f_n(x) \in \mathbb{R}\right\}$

For the limit of a sequence of functions we have a simpler domain then for the limit of partial functions.

Lemma 18.89. *Let X be a set and $\{f_n: X \rightarrow \mathbb{R}\}_{n \in \mathbb{N}}$ a sequence of functions [so $\forall n \in \mathbb{N}$ $\operatorname{dom}(f_n) = X$) then*

1. $\operatorname{dom}\left(\lim_{n \rightarrow \infty} f_n\right) = \{x \in X | \{x_n\}_{n \in \mathbb{N}} \text{ has a limit in } \mathbb{R}\}$
2. $\forall x \in \operatorname{dom}\left(\lim_{n \rightarrow \infty} f_n\right)$ $\left(\lim_{n \rightarrow \infty} f_n\right)(x)$ *is the limit of* $\{f_i(x)\}_{i \in \mathbb{N}}$
3. *If* $\forall n \in \mathbb{N}$ *we have* $f_n \leqslant f_{n+1}$ *then* $\left(\lim_{n \rightarrow \infty} f_i\right)(x) = \sup(\{f_i(x) | i \in \mathbb{N}\})$

Proof.

1. First using 18.88 we have that $\operatorname{dom}\left(\lim_{n - < \infty} f_n\right) = \big\{x \in \bigcup_{n \in \mathbb{N}} \left(\bigcap_{i \in \{n, \ldots, \infty\}} \operatorname{dom}(f_i)\right) | \exists k \in \mathbb{N} \text{ such that } x \in \bigcap_{i \in \{k, \ldots, \infty\}} \operatorname{dom}(f_i)$ and $\{f_n(x)\}_{n \in \{k, \ldots \infty\}} \text{ has a limit in } \mathbb{R}\big\}$. Now $\forall n \in \mathbb{N}$ we have $\operatorname{dom}(f_n) = X$ so that $\bigcup_{n \in \mathbb{N}} \left(\bigcap_{i \in \{n, \ldots, \infty\}} \operatorname{dom}(f_i)\right) = \bigcup_{n \in \mathbb{N}} X = X$. Hence we have

$$
\begin{aligned}
x \in \operatorname{dom}\left(\lim_{n \rightarrow \infty} f_n\right) &\Leftrightarrow x \in X \wedge \exists k \in \mathbb{N} \text{ with } x \in \bigcap_{n \in \{k, \ldots, \infty\}} \operatorname{dom}(f_n) \wedge \\
&\qquad \{f_n\}_{n \in \{k, \ldots, \infty\}} \text{ has a limit} \\
&\Leftrightarrow x \in X \wedge \exists k \in \mathbb{N} \text{ with } x \in X \wedge \\
&\qquad \{f_n\}_{n \in \{k, \ldots, \infty\}} \text{ has a limit} \\
&\underset{12.304}{\Leftrightarrow} x \in X \text{ and } \{f_n\}_{n \in \{1, \ldots, \infty\}} \text{ has a limit}
\end{aligned}
$$

   which proves that $\operatorname{dom}\left(\lim_{n - < \infty} f_n\right) = \{x \in X | \{f_n(x)\}_{n \in \mathbb{N}} \text{ has a limit in } \mathbb{R}\}$.

2. This follows from 18.88 and 12.304.

3. This follows from 12.354 and the fact that $\forall n \in \mathbb{N}\, f_n(x) \leqslant f_{n+1}(x)$ □

Proposition 18.90. *Let X be a set, $C \subseteq X$ and $f: X \rightarrow \mathbb{R}$ a partial function then $|f_{|C}| = |f|_{|C}$*

Proof. First $\operatorname{dom}(|f_{|C}|) = \operatorname{dom}(f_{|C}) = C \bigcap \operatorname{dom}(f) = C \bigcap \operatorname{dom}(|f|) = \operatorname{dom}(|f|_{|C})$ and $\forall x \in \operatorname{dom}(f_{|C})$ we have $(|f|_{|C})(x) = |f|(x) = |f(x)| = |(f_{|C})(x)| = |f_{|C}|(x)$ □

Definition 18.91. *Let $\langle X, \mathcal{A}\rangle$, $\langle Y, \mathcal{B}\rangle$ be measurable spaces, then the partial function $f: X \rightarrow Y$ is $\mathcal{A}/\mathcal{B}$-**measurable** if $\forall B \in \mathcal{B}$ we have that $f^{-1}(B) \in \mathcal{A}_{|\operatorname{dom}(f)}$ [here $\mathcal{A}_{|\operatorname{dom}(f)} = \{\operatorname{dom}(f) \bigcap A | A \in \mathcal{A}\}$ is the sub-algebra induced by $\operatorname{dom}(f)$ (see 18.9)].*

Note 18.92. If $\operatorname{dom}(f) = X$ then $\mathcal{A}_{|\operatorname{dom}(f)} = \mathcal{A}$ so in this case the **function** f: $X \rightarrow Y$ is $\mathcal{A}/\mathcal{B}$-measurable if $\forall B \in \mathcal{B}$ we have that $f^{-1}(B) \in \mathcal{A}$.

As we have to deal a lot with partial functions to the real numbers we introduce the following notations.

Notation 18.93.

1. *If $\langle X, \mathcal{A}\rangle$ is a measurable space X then the partial function $f: X \rightarrow \mathbb{R}$ is $\mathcal{A}$-measurable iff f is $\mathcal{A}/\mathcal{B}[\mathbb{R}]$-measurable.*
2. *If X is a set, $\mathcal{A} \subseteq \mathcal{P}(X)$, then the partial function $f: X \rightarrow \mathbb{R}$ is Borel measurable iff f is $\sigma[\mathcal{A}]/\mathcal{B}[\mathbb{R}]$-measurable.*
3. *If $\langle X, \mathcal{T}\rangle$ is a topological space, then the partial function $f: D \rightarrow \mathbb{R}$ is Borel measurable iff f is $\mathcal{B}[X, \mathcal{T}] : \mathcal{B}[\mathbb{R}]$-measurable.*
4. *The partial function $f: \mathbb{R} \rightarrow \mathbb{R}$ is Borel measurable iff f is $\mathcal{B}[\mathbb{R}]/\mathcal{B}[\mathbb{R}]$-measurable.*
5. *The partial function $f: \mathbb{R} \rightarrow \mathbb{R}$ is Lebesgue measurable iff f is $\mathcal{L}/\mathcal{B}[\mathbb{R}]$-measurable*
6. *The partial function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is Borel measurable iff f is $\mathcal{B}[\mathbb{R}^n]/\mathcal{B}[\mathbb{R}]$-measurable.*
7. *The partial function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is Lebesgue measurable iff f is $\mathcal{L}^n/\mathcal{B}[\mathbb{R}]$-measurable.*

For Borel algebras (see 18.11) we can use the generator of the Borel algebra to prove if a function is measurable.

Theorem 18.94. *Let $\langle X, \mathcal{A}\rangle$ be a measurable space X, Y a set, $\mathcal{B} \subseteq \mathcal{P}(Y)$ and $f: X \rightarrow Y$ a partial function then f is $\mathcal{A}/\sigma[\mathcal{B}]$-measurable if and only if $\forall B \in \mathcal{B}$ $f^{-1}(B) \in \mathcal{A}_{|\mathrm{dom}(f)}$*

Proof.

⇒. If f is $\mathcal{A}/\sigma[\mathcal{B}]$-measurable then $\forall B \in \mathcal{B} \subseteq \sigma[\mathcal{B}]$ we have that $f^{-1}(B) \in \mathcal{A}_{|\mathrm{dom}(f)}$

⇐. Assume that $\forall B \in \mathcal{B}$ $f^{-1}(B) \in \mathcal{A}_{|\mathrm{dom}(f)}$. Define $\mathcal{C} = \{B \in \mathcal{P}(Y) | f^{-1}(B) \in \mathcal{A}_{|\mathrm{dom}(f)}\}$ then by assumption we have

$$\mathcal{B} \subseteq \mathcal{C} \tag{18.150}$$

As $f^{-1}(\emptyset) = \emptyset \in \mathcal{A}_{|\mathrm{dom}(f)}$ we have

$$\emptyset \in \mathcal{C} \tag{18.151}$$

If $B \in \mathcal{C}$ then by definition $f^{-1}(B) \in \mathcal{A}_{|\mathrm{dom}(f)}$, as $f^{-1}(Y \setminus B) \underset{2.54\,(4)}{=} \mathrm{dom}(f) \setminus f^{-1}(B) \in \mathcal{A}_{|\mathrm{dom}(f)}$ [as $\mathcal{A}_{|\mathrm{dom}(f)}$ is a σ-algebra on $\mathrm{dom}(f)$] proving that

$$\forall B \in \mathcal{C} \text{ we have } Y \setminus B \in \mathcal{C} \tag{18.152}$$

If $\{B_i\}_{i \in \mathbb{N}} \subseteq \mathcal{C}$ then $\forall i \in \mathbb{N}$ we have $f^{-1}(B) \in \mathcal{A}_{|\mathrm{dom}(f)}$. Now $f^{-1}(\bigcup_{i \in \mathbb{N}} B_i) \underset{2.58}{=} \bigcup_{i \in \mathbb{N}} f^{-1}(B_i) \in \mathcal{A}_{|\mathrm{dom}(f)}$ proving that

$$\forall \{B_i\}_{i \in \mathbb{N}} \subseteq \mathcal{C} \text{ we have } \bigcup_{i \in \mathbb{N}} B_i \in \mathcal{C} \tag{18.153}$$

To summarize (18.150),(18.151),(18.152) and (18.153) proves that $\mathcal{C}$ is a σ-algebra containing $\mathcal{B}$, hence $\sigma[\mathcal{B}] \subseteq \mathcal{C}$ or in other words

$$\forall B \in \sigma[\mathcal{B}] \text{ we have } f^{-1}(B) \in \mathcal{A}_{|\mathrm{dom}(f)} \qquad \square$$

Continuous functions are automatically measurable as the following proposition shows.

Proposition 18.95. *Let $\langle X, \mathcal{T}_X\rangle$, $\langle Y, \mathcal{T}_Y\rangle$ be two topological spaces, X and f: $X \rightarrow Y$ is a partial function so that f: $\mathrm{dom}(f) \rightarrow Y$ is a continuous function (using the subspace topology on $\mathrm{dom}(f)$) then f is $\mathcal{B}[X, \mathcal{T}_X] / \mathcal{B}[Y, \mathcal{T}_Y]$-measurable.*

Proof. As f is continuous using the subspace topology we have $\forall V \in \mathcal{T}$ that $f^{-1}(V) \in (\mathcal{T}_X)_{|\mathrm{dom}(f)}$, so there exists a $U \in \mathcal{T}_X$ such that $f^{-1}(V) = U \bigcap \mathrm{dom}(f)$. As $\mathcal{T}_X \subseteq \sigma[\mathcal{T}_X] = \mathcal{B}[X, \mathcal{T}_X]$ it follows that $f^{-1}(V) \in \mathcal{B}[X, \mathcal{T}_X]_{|\mathrm{dom}(f)}$. Finally using the fact that $\mathcal{B}[Y, \mathcal{T}_Y] = \sigma[\mathcal{T}_Y]$ it follows using 18.94 that f is $\mathcal{B}[X, \mathcal{T}_X] / \mathcal{B}[Y, \mathcal{T}_Y]$-measurable. $\square$

The composition of measurable mappings is measurable as is expressed in the following proposition.

Proposition 18.96. *Let $\langle X, \mathcal{A}\rangle$, $\langle Y, \mathcal{B}\rangle$ and $\langle Z, \mathcal{C}\rangle$ be measurable spaces, X, Y, f: $X \rightarrow Y$ a $\mathcal{A}/\mathcal{B}$-measurable partial function , g: $Y \rightarrow Z$ a $\mathcal{B}/\mathcal{C}$-measurable partial function then the partial function $g \circ f$: $X \rightarrow Z$ is a $\mathcal{A}/\mathcal{C}$-measurable partial function.*

Proof. Let $C \in \mathcal{C}$ then as g is $\mathcal{B}/\mathcal{C}$-measurable we have that $g^{-1}(C) \in \mathcal{B}_{|\mathrm{dom}(g)}$ so there exists a $B' \in \mathcal{B}$ such that $g^{-1}(C) = B' \bigcap \mathrm{dom}(g)$. Then $f^{-1}(g^{-1}(C)) = f^{-1}(B' \bigcap \mathrm{dom}(g)) \underset{2.58}{=} f^{-1}(B') \bigcap f^{-1}(\mathrm{dom}(g))$ proving that

$$f^{-1}(g^{-1}(C)) = f^{-1}(B') \bigcap f^{-1}(\mathrm{dom}(g)) \tag{18.154}$$

As f is $\mathcal{A}/\mathcal{B}$-measurable function we have that $f^{-1}(B') \in \mathcal{A}_{|\mathrm{dom}(f)}$ so that $\exists A' \in \mathcal{A}$ such that $f^{-1}(B') = A' \bigcap \mathrm{dom}(f)$ hence $f^{-1}(g^{-1}(C)) \underset{(18.154)}{=} (A' \bigcap \mathrm{dom}(f)) \bigcap f^{-1}(\mathrm{dom}(g)) = A' \bigcap (\mathrm{dom}(f) \bigcap f^{-1}(\mathrm{dom}(g))) \in \mathcal{A}_{|\mathrm{dom}(f) \bigcap f^{-1}(\mathrm{dom}(g))}$ which proves that $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C)) \in \mathcal{A}_{|\mathrm{dom}(f) \bigcap f^{-1}(\mathrm{dom}(g))}$. Finally as by 2.8 $\mathrm{dom}(g \circ f) = \mathrm{dom}(f) \bigcap f^{-1}(\mathrm{dom}(g))$ it follows that $g \circ f$ is $\mathcal{A}/\mathcal{C}$-measurable. $\square$

A important case of measurable functions in integration theory is the case were $\langle Y, \mathcal{B}\rangle = \langle \mathbb{R}, \mathcal{B}[\mathbb{R}]\rangle$, we have then the following equivalences for measurable functions. First we have some simplifying definitions

Definition 18.97. *Let f: $X \rightarrow \mathbb{R}$ is a partial function then for $a, b \in \mathbb{R}$*

1. $\{f < a\} = \{x \in \mathrm{dom}(f) | f(x) < a\} \subseteq \mathrm{dom}(f)$
2. $\{f \leqslant a\} = \{x \in \mathrm{dom}(f) | f(x) \leqslant a\} \subseteq \mathrm{dom}(f)$
3. $\{a < f\} = \{x \in \mathrm{dom}(f) | a < f(x)\} \subseteq \mathrm{dom}(f)$
4. $\{a \leqslant f\} = \{x \in \mathrm{dom}(f) | a \leqslant f(x)\} \subseteq \mathrm{dom}(f)$

Theorem 18.98. *Let $\langle X,\mathcal{A}\rangle$ be a measurable space and $f\colon X\to\mathbb{R}$ a partial function then the following are equivalent*

1. *f is $\mathcal{A}/\mathcal{B}[\mathbb{R}]$-measurable*
2. *$\forall a\in\mathbb{R}\;\{f<a\}\in\mathcal{A}_{|\mathrm{dom}(f)}$*
3. *$\forall a\in\mathbb{R}\;\{f\leqslant a\}\in\mathcal{A}_{|\mathrm{dom}(f)}$*
4. *$\forall a\in\mathbb{R}\;\{a<f\}\in\mathcal{A}_{|\mathrm{dom}(f)}$*
5. *$\forall a\in\mathbb{R}\;\{a\leqslant f\}\in\mathcal{A}_{|\mathrm{dom}(f)}$*

Proof.

1 ⇔ 2. Let $a\in\mathbb{R}$ then

$$\begin{aligned} x\in f^{-1}(]-\infty,a[) &\Leftrightarrow x\in\mathrm{dom}(f)\wedge f(x)\in]-\infty,a[\\ &\Leftrightarrow x\in\mathrm{dom}(f)\wedge f(x)<a\\ &\Leftrightarrow x\in\{f<a\}\end{aligned}$$

proving that $f^{-1}(]-\infty,a[)=\{f<a\}$, as $\mathcal{B}[\mathbb{R}]\underset{18.62}{=}\sigma[\{]-\infty,a[|a\in\mathbb{R}\}]$, we have by 18.94 that f is $\mathcal{A}/\mathcal{B}[\mathbb{R}]$-measurable if and only if $\{f<a\}\in\mathcal{A}_{|\mathrm{Adom}(f)}$.

1 ⇔ 3. Let $a\in\mathbb{R}$ then

$$\begin{aligned} x\in f^{-1}(]-\infty,a]) &\Leftrightarrow f(x)\in]-\infty,a]\\ &\Leftrightarrow f(x)\leqslant a\\ &\Leftrightarrow x\in\{f\leqslant a\}\end{aligned}$$

proving that $f^{-1}(]-\infty,a])=\{f\leqslant a\}$, as $\mathcal{B}[\mathbb{R}]\underset{18.62}{=}\sigma[\{]-\infty,a]|a\in\mathbb{R}\}]$, we have by 18.94 that f is $\mathcal{A}/\mathcal{B}[\mathbb{R}]$-measurable if and only if $\{f\leqslant a\}\in\mathcal{A}_{|\mathrm{dom}(f)}$

1 ⇔ 4. Let $a\in\mathbb{R}$ then

$$\begin{aligned} x\in f^{-1}(]a,\infty[) &\Leftrightarrow f(x)\in]a,\infty[\\ &\Leftrightarrow a<f(x)\\ &\Leftrightarrow x\in\{a<f\}\end{aligned}$$

proving that $f^{-1}(]a,\infty[)=\{a<f\}$, as $\mathcal{B}[\mathbb{R}]\underset{18.62}{=}\sigma[\{]a,\infty[|a\in\mathbb{R}\}]$, we have by 18.94 that f is $\mathcal{A}/\mathcal{B}[\mathbb{R}]$-measurable if and only if $\{a<f\}\in\mathcal{A}_{|\mathrm{dom}(f)}$

4 ⇒ 5. Let $a\in\mathbb{R}$ then

$$\begin{aligned} x\in f^{-1}([a,\infty[) &\Leftrightarrow f(x)\in[a,\infty[\\ &\Leftrightarrow a\leqslant f(x)\\ &\Leftrightarrow x\in\{a\leqslant f\}\end{aligned}$$

proving that $f^{-1}([a,\infty[)=\{a\leqslant f\}$, as $\mathcal{B}[\mathbb{R}]\underset{18.62}{=}\sigma[\{[a,\infty[|a\in\mathbb{R}\}]$, we have by 18.94 that f is $\mathcal{A}/\mathcal{B}[\mathbb{R}]$-measurable if and only if $\{a\leqslant f\}\in\mathcal{A}_{|\mathrm{dom}(f)}$ □

Example 18.99. The function $|.|\colon\mathbb{R}\to\mathbb{R}$ is $\mathcal{B}[\mathbb{R}]$-measurable

Proof. Let $a\in\mathbb{R}$ then for $\{|.|\leqslant a\}$ we have the following cases to consider as $\mathrm{dom}(|.|)=\mathbb{R}$

$0\leqslant a$. then

$$\begin{aligned} x\in\{|.|\leqslant a\} &\Leftrightarrow |x|\leqslant a\\ &\Leftrightarrow x\leqslant a\vee -x\leqslant a\\ &\Leftrightarrow x\leqslant a\vee -a\leqslant x\\ &\Leftrightarrow x\in\,]-\infty,a]\vee x\in[-a,\infty[\\ &\Leftrightarrow x\in\,]-\infty,a]\bigcup[-a,\infty[\end{aligned}$$

proving that $\{|.|\leqslant a\}=\,]-\infty,a]\bigcup[-a,\infty[$. As by 18.62 $]-\infty,a],[-a,\infty[\,\in\mathcal{B}[\mathbb{R}]$ we have that

$$\{|.|\leqslant a\}\in\mathcal{B}[\mathbb{R}]$$

$a<0$. then as $\forall x\in\mathbb{R}$ we have that $0\leqslant|x|$ it follows that

$$\{|.|\leqslant a\}=\emptyset\in\mathcal{B}[\mathbb{R}]$$

□

Lemma 18.100. *Let $\langle X,\mathcal{A}\rangle$ be a measurable space, $E\subseteq X$ and $f:X\rightarrow\mathbb{R}$ a $\mathcal{A}$-measurable function then $\forall a\in\mathbb{R}$ we have that*

1. $\{f<a\}\bigcap E\in\mathcal{A}_{|\mathrm{dom}(f)\bigcap E}$
2. $\{f\leqslant a\}\bigcap E\in\mathcal{A}_{|\mathrm{dom}(f)\bigcap E}$
3. $\{a<f\}\bigcap E\in\mathcal{A}_{|\mathrm{dom}(f)\bigcap E}$
4. $\{a\leqslant f\}\bigcap E\in\mathcal{A}_{|\mathrm{dom}(f)\bigcap E}$

Proof.

1. As f is $\mathcal{A}$-measurable we have that $\{f<a\}\in\mathcal{A}_{|f(x)}$ so there exists a $A\in\mathcal{A}$ such that $\{f<a\}=A\bigcap\mathrm{dom}(f)$. Hence $\{f<a\}\bigcap E=(A\bigcap\mathrm{dom}(f))\bigcap E=A\bigcap(\mathrm{dom}(f)\bigcap A)\in\mathcal{A}_{|\mathrm{dom}(f)\bigcap E}$
2. As f is $\mathcal{A}$-measurable we have that $\{f\leqslant a\}\in\mathcal{A}_{|\mathrm{dom}(f)}$ so there exists a $A\in\mathcal{A}$ such that $\{f\leqslant a\}=A\bigcap\mathrm{dom}(f)$. Hence $\{f\leqslant a\}\bigcap E=(A\bigcap\mathrm{dom}(f))\bigcap E=A\bigcap(\mathrm{dom}(f)\bigcap A)\in\mathcal{A}_{|\mathrm{dom}(f)\bigcap A}$
3. As f is $\mathcal{A}$-measurable we have that $\{a<f\}\in\mathcal{A}_{|\mathrm{dom}(f)}$ so there exists a $A\in\mathcal{A}$ such that $\{a<f\}=A\bigcap\mathrm{dom}(f)$. Hence $\{a<f\}\bigcap E=(A\bigcap\mathrm{dom}(f))\bigcap E=A\bigcap(\mathrm{dom}(f)\bigcap A)\in\mathcal{A}_{|\mathrm{dom}(f)\bigcap A}$
4. As f is $\mathcal{A}$-measurable we have that $\{a\leqslant f\}\in\mathcal{A}_{|\mathrm{dom}(f)}$ so there exists a $A\in\mathcal{A}$ such that $\{a\leqslant f\}=A\bigcap\mathrm{dom}(f)$. Hence $\{a\leqslant f\}\bigcap E=(A\bigcap\mathrm{dom}(f))\bigcap E=A\bigcap(\mathrm{dom}(f)\bigcap A)\in\mathcal{A}_{|\mathrm{dom}(f)\bigcap A}$ □

Definition 18.101. *A partial function $f:\mathbb{R}\rightarrow\mathbb{R}$ is strict monotonic if either*

1. $\forall x,y\in\mathrm{dom}(f)$ *with* $x<y$ *we have* $f(x)<f(y)$ *(f is strict increasing)*

2. *$\forall x, y \in \mathrm{dom}(f)$ with $x < y$ we have $f(y) < f(x)$ (f is strict decreasing)*

Proposition 18.102. *Let $n \in \mathbb{N}$ and $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a partial function then*

1. *If f is Borel measurable then f is Lebesgue measurable.*
2. *If f is continuous (using the subspace topology on $\mathrm{dom}(f)$) then f is Borel measurable.*
3. *If $n = 1$ and f is strict monotonic then f is Borel measurable.*

Proof.

1. As f is Borel measurable we have using 18.98 that $\forall a \in \mathbb{R}$ $\{f < a\} \in \mathcal{B}[\mathbb{R}^n]_{|\mathrm{dom}(f)} \subseteq \mathcal{L}^n_{|\mathrm{dom}(f)}$ $[\mathcal{B}[\mathbb{R}^n]_{|\mathrm{dom}(f)}\{B \bigcap \mathrm{dom}(f) | B \in \mathcal{B}[\mathbb{R}^n]\} \subseteq_{18.62} \{B \bigcap \mathrm{dom}(f) | N \in \mathcal{L}^n\} = \mathcal{L}^n_{|\mathrm{dom}(f)}]$ hence using 18.98 again we have that f is Lebesgue measurable.
2. This follows from 18.95.
3. As f is monotonic we have that either

   ***f* is non decreasing.** So $\forall x, y \in \mathrm{dom}(f)$ with $x < y$ we have $f(x) < f(y)$. Let now $a \in \mathbb{R}$ then for $\{f < a\}$ we have the following possibilities:

   $\{f < a\} = \emptyset$. then $\{f < a\} = \emptyset \in \mathcal{B}[\mathbb{R}]_{|\mathrm{dom}(f)}$

   $\{f < a\} = \mathrm{dom}(f)$. then $\{f < a\} = \mathbb{R} \bigcap \mathrm{dom}(f) \in \mathcal{B}[\mathbb{R}]_{|\mathrm{dom}(f)}$

   $\emptyset \neq \{f < a\} \neq \mathrm{dom}(f)$. then as $\{f < a\} \subseteq \mathrm{dom}(f)$ there exists a $x \in \mathrm{dom}(f)$ such that $x \notin \{f < a\}$ hence $a \leqslant f(x)$. Let $y \in \{f < a\} \Rightarrow f(y) < a$ and assume that $x < y$ then $a \leqslant f(x) < f(y)$ contradicting $f(y) < a$. Hence x is a upper bound of $\{f < a\} \neq \emptyset$ so that $s = \sup(\{f < a\})$ exists. We must consider now two cases for s

   $s \in \{f < a\}$. As s is a upper bound of $\{f < a\}$ it follows that

   $$\{f < a\} \subseteq]-\infty, s] \bigcap \mathrm{dom}(f) \tag{18.155}$$

   If $x \in]-\infty, s] \bigcap \mathrm{dom}(f)$ then $x \leqslant s \Rightarrow f(x) \leqslant f(s) < a$ so $x \in \{f < a\}$ which together with (18.155) proves that $\{f < a\} =]-\infty, s] \bigcap D \in \sigma[\{]-\infty, a] | a \in \mathbb{R}\}]_{|\mathrm{dom}(f)} \underset{18.62}{=} \mathcal{B}[\mathbb{R}]_{|\mathrm{dom}(f)}$ proving

   $$\{f < a\} \in \mathcal{B}[\mathbb{R}]_{|\mathrm{dom}(f)}$$

   $s \notin \{f < a\}$. As s is a upper bound of $\{f < a\}$it follows that

   $$\{f < a\} \subseteq]-\infty, s[\bigcap \mathrm{dom}(f) \tag{18.156}$$

If $x \in]-\infty, s[\bigcap \text{dom}(f)$ then $x < s$ so by the definition of a supremum there exists a $y \in \{f < a\} \Rightarrow f(y) < a$ such that $x < y \leqslant s$ then $f(x) < f(y) < a$ proving that $x \in \{f < a\}$. Together with (18.156) this proves that $\{f < a\} =]-\infty, s[\bigcap D \in \sigma[\{]-\infty, a[|a \in \mathbb{R}\}]_{|\text{dom}(f)} \underset{18.62}{=} \mathcal{B}[\mathbb{R}]_{|\text{dom}(f)}$ proving

$$\{f < a\} \in \mathcal{B}[\mathbb{R}]_{|\text{dom}(f)}$$

So in all cases we have that $\{f < a\} \in \mathcal{B}[\mathbb{R}]_{|\text{dom}(f)}$ proving as $a \in \mathbb{R}$ by 18.98 that f is measurable.

***f* is non increasing.** So $\forall x, y \in \text{dom}(f)$ with $x < y$ we have $f(y) < f(x)$. Let now $a \in \mathbb{R}$ then for $\{a < f\}$ we have the following possibilities:

$\boldsymbol{\{a < f\} = \emptyset}$**.** then $\{a < f\} = \emptyset \in \mathcal{B}[\mathbb{R}]_{|\text{dom}(f)}$

$\boldsymbol{\{a < f\} = \text{dom}(f)}$**.** then $\{a < f\} = \mathbb{R}\bigcap D \in \mathcal{B}[\mathbb{R}]_{|\text{dom}(f)}$

$\boldsymbol{\emptyset \neq \{a < f\} \neq \text{dom}(f)}$**.** then as $\{a < f\} \subseteq \text{dom}(f)$ there exists a $x \in \text{dom}(f)$ such that $x \notin \{a < f\}$ hence $f(x) \leqslant a$. Let $y \in \{a < f\} \Rightarrow a < f(y)$ and assume that $x < y$ then $a < f(y) < f(x) \leqslant a$ a contradiction. So $\emptyset \neq \{a < f\}$ is bounded above by x and thus $s = \sup(\{a < f\})$ exists. We must consider the following cases for s:

$\boldsymbol{s \in \{a < f\}}$**.** As s is a upper bound of $\{a < f\}$ we have

$$\{a < f\} \subseteq]-\infty, s]\bigcap \text{dom}(f) \tag{18.157}$$

If $x \in]-\infty, s]\bigcap \text{dom}(f)$ then $x \leqslant s \Rightarrow a < f(s) \leqslant f(x)$ proving that $x \in \{a < f\}$ which together with (18.157) proves that $\{a < f\} =]-\infty, s]\bigcap D \in \sigma[\{]-\infty, a]|a \in \mathbb{R}\}]_{|\text{dom}(f)} \underset{18.62}{=} \mathcal{B}[\mathbb{R}]_{|\text{dom}(f)}$ proving

$$\{a < f\} \in \mathcal{B}[\mathbb{R}]_{|\text{dom}(f)}$$

$\boldsymbol{s \notin \{a < f\}}$**.** As s is a upper bound of $\{a < f\}$ we have

$$\{a < f\} \subseteq]-\infty, s[\bigcap \text{dom}(f) \tag{18.158}$$

Let $x \in]-\infty, s[\bigcap \text{dom}(f)$ then $x < s$ then using the definition of a supremum there exists a $y \in \{a < f\} \Rightarrow a < f(y)$ such that $x < y \leqslant s$ then $a < f(y) < f(x)$ proving that $x \in \{a < f\}$. Using this together with (18.158) proves that $\{a < f\} =]-\infty, s[\bigcap \text{dom}(f) \in \sigma[\{]-\infty, a[|a \in \mathbb{R}\}]_{|\text{dom}(f)} \underset{18.62}{=} \boldsymbol{B}[\mathbb{R}]_{|\text{dom}(f)}$. Hence

$$\{a < f\} \in \mathcal{B}[\mathbb{R}]_{|\text{dom}(f)}$$

So in all cases we have that $\{f < a\} \in \mathcal{B}[\mathbb{R}]_{|\text{dom}(f)}$ proving as $a \in \mathbb{R}$ by 18.98 that f is measurable. □

The following theorem shows the operations that we can do on measurable real valued functions.

Theorem 18.103. *Let $\langle X, \mathcal{A}\rangle$ be a measurable set, X and $f: X \to \mathbb{R}$, $g: X \to \mathbb{R}$ partial functions then we have:*

1. *If f is a constant function [$\forall x \in \operatorname{dom}(f)$ we have $f(x) = c$ where $c \in \mathbb{R}$] then f is $\mathcal{A}$-measurable.*
2. *If f, g are $\mathcal{A}$-measurable functions then $f + g: X \to \mathbb{R}$ [see 18.88] is $\mathcal{A}$-measurable.*
3. *If $n \in \mathbb{N}$ and $\{f_i: X \to \mathbb{R}\}_{i \in \{1,\ldots,n\}}$ a finite family of $\mathcal{A}$-measurable partial functions then $(\sum_{i=1}^{n} f_i): X \to \mathbb{R}$ [see 18.88] is $\mathcal{A}$-measurable.*
4. *If f is a $\mathcal{A}$-measurable partial function, $c \in \mathbb{R}$ then $c \cdot f: X \to \mathbb{R}$ [see 18.88] is $\mathcal{A}$-measurable.*
5. *If f, g are $\mathcal{A}$-measurable partial functions then $f \cdot g: X \to \mathbb{R}$ [see 18.88] is $\mathcal{A}$-measurable.*
6. *If f, g are $\mathcal{A}$-measurable partial functions then $f / g: X \to \mathbb{R}$ [see 18.88] is $\mathcal{A}$-measurable.*
7. *If f is a $\mathcal{A}$-measurable partial function and $C \subseteq X$ then $f_{|C}: X \to \mathbb{R}$ is a $\mathcal{A}$-measurable partial function [note that $\operatorname{dom}(f) = C \bigcap \operatorname{dom}(f)$]*
8. *If f is a $\mathcal{A}$-measurable partial function then $|f|: X \to \mathbb{R}$ [see 18.88] is a $\mathcal{A}$-measurable function.*

Proof.

1. Let $a \in \mathbb{R}$ then we have for c the following possibilities

   $a \leqslant c$. then if $x \in \{f < a\} \Rightarrow x \in \operatorname{dom}(f) \wedge f(x) < a \underset{f(x)=c}{\Rightarrow} c < a \leqslant c$ a contradiction, so $\{f < a\} = \emptyset \in \mathcal{A}_{|\operatorname{dom}(f)}$

   $c < a$. then

   $$\begin{array}{rcl} x \in \{f < a\} & \Leftrightarrow & x \in \operatorname{dom}(f) \wedge f(x) < a \\ & \Leftrightarrow & x \in \operatorname{dom}(f) \wedge c < a \\ & \underset{c<a \text{ assumed}}{\Leftrightarrow} & x \in \operatorname{dom}(f) \end{array}$$

   proving that $\{f < a\} = \operatorname{dom}(f) \in \mathcal{A}_{|\operatorname{dom}(f)}$.

   So in all cases we have $\{f < A\} \in \mathcal{A}_{|\operatorname{dom}(f)}$ which by 18.98 proves that f is $\mathcal{A}$-measurable.

2. Let $a \in \mathbb{R}$ and define $Q_a = \{(q, q') \in \mathbb{Q} \times \mathbb{Q} | q + q' < a\} \subseteq \mathbb{Q} \times \mathbb{Q}$ then using 9.38,5.69 and 5.66 it follows that

   $$Q_a \text{ is countable} \tag{18.159}$$

   Given $q \in \mathbb{Q}$ we have as f, g are $\mathcal{A}$-measurable that $\{f < q\} \in \mathcal{A}_{|\operatorname{dom}(f)}$ and $\{g < q\} \in \mathcal{A}_{|\operatorname{dom}(g)}$. So there exists $F_q, G_q \in \mathcal{A}$ such that

   $$\{f < q\} = F_q \bigcap \operatorname{dom}(f) \text{ and } \{g < q\} = G_q \bigcap \operatorname{dom}(g) \tag{18.160}$$

For each $(q,q')\in Q_a$ define $E_{q,q'}=\{x\in D\bigcap E|f(x)<q\wedge g(x)<q'\}$ then

$$\begin{aligned}
x\in E_{q,q'} &\Leftrightarrow x\in \mathrm{dom}(f)\bigcap \mathrm{dom}(g)\wedge f(x)<q\wedge g(x)<q'\\
&\Leftrightarrow \big(x\in \mathrm{dom}(f)\bigcap \mathrm{dom}(g)\big)\wedge (x\in \mathrm{dom}(f)\wedge f(x)<q)\wedge (x\in \\
&\quad \mathrm{dom}(g)\wedge g(x)<q')\\
&\Leftrightarrow x\in \mathrm{dom}(f)\bigcap \mathrm{dom}(g)\wedge x\in\{f<q\}\wedge x\in\{g<q'\}\\
&\Leftrightarrow x\in \mathrm{dom}(f)\bigcap \mathrm{dom}(g)\wedge x\in F_q\bigcap \mathrm{dom}(f)\wedge x\in G_{q'}\bigcap \mathrm{dom}(g)\\
&\Leftrightarrow x\in\big(\mathrm{dom}(f)\bigcap \mathrm{dom}(g)\big)\bigcap\big(F_q\bigcap \mathrm{dom}(f)\big)\bigcap\big(G_{q'}\bigcap \mathrm{dom}(g)\big)\\
&\Leftrightarrow x\in\big(\mathrm{dom}(f)\bigcap \mathrm{dom}(g)\big)\bigcap F_q\bigcap G_{q'}
\end{aligned}$$

proving that

$$\forall (q,\ q')\in Q_a \text{ we have } E_{q,q'}=\big(\mathrm{dom}(f)\bigcap\ \mathrm{dom}(g)\big)\bigcap\ F_q\bigcap\ G_{q'}\in \mathcal{A}_{|D\bigcap E} \tag{18.161}$$

Now if for $x\in\mathrm{dom}(f)\bigcap\mathrm{dom}(g)$ we have $(f+g)(x)=f(x)+g(x)<a$ then $f(x)<a-g(x)$ so using the density theorem of the real numbers (see 9.57) we have

$$\exists q\in Q \,\text{such that}\, f(x)<q<a-g(x) \tag{18.162}$$

Further from $f(x)+g(x)<a$ we have also $g(x)<_{(18.162)} a-q$ and using 9.57 again we have

$$\exists q'\in Q \,\text{such that}\, g(x)<q'<a-q \tag{18.163}$$

From the (18.163) we have then that $q+q'<a$ hence $(q,q')\in Q_a$ and as $f(x)<q$, $g(x)<q'$ it follows that $x\in E_{q,q'}$. So if $x\in\{f+g<a\}=\{x\in \mathrm{dom}(f)\bigcap\mathrm{dom}(g)|(f+g)(x)<a\}$ we have that $\exists (q,q')\in Q_a$ such that $(q, q')\in E_{q,q'}$ proving that

$$\{f+g<a\}\subseteq \bigcup_{(q,q')\in Q_a} E_{q,q'} \tag{18.164}$$

Further if $x\in\bigcup_{(q,q')\in Q_a}E_{q,q'}$ there exists a $(q,q')\in Q_a$ such that $x\in E_{q,q'}$ hence $x\in\mathrm{dom}(f)\bigcap\mathrm{dom}(g)$ and $f(x)<q\wedge g(x)<q'\Rightarrow (f+g)(x)=f(x)+g(x)<q+q'<a$ [as $(q,q')\in Q_a$] proving that $x\in\{f+g<a\}$ and thus $\bigcup_{(q,q')\in Q_a}E_{q,q'}\subseteq\{f+g<a\}$. Using (18.164) it follows then that

$$\{f+g<a\}=\bigcup_{(q,q')\in Q_a} E_{q,q'} \tag{18.165}$$

As $\bigcup_{(q,q')\in Q_a}E_{q,q'}$ is a countable union of sets in $\mathcal{A}_{|\mathrm{dom}(f)\bigcap\mathrm{dom}(g)}$ (see (18.161)) it follows from 18.6 that $\bigcup_{(q,q')\in Q_a}E_{q,q'}\in\mathcal{A}_{|\mathrm{dom}(f)\bigcap\mathrm{dom}(g)}$ or using (18.165) that $\{f+g<a\}\in\mathcal{A}_{|\mathrm{dom}(f)\bigcap\mathrm{dom}(g)}$. So as $a\in\mathbb{R}$ was chosen arbitrary we have by 18.98

$$f+g\,\text{is}\,\mathcal{A}\text{-measurable}$$

3. We prove this by induction, so let $\mathcal{N}=\{n\in\mathbb{N}|\forall\{f_i: X\rightarrow\mathbb{R}\}_{i\in\{1,\ldots,n\}}$ $\mathcal{A}$-measurable partial functions we have that $\sum_{i=1}^{n} f_i$ is $\mathcal{A}$-measurable$\}$ then

   $1\in\mathcal{N}$. this is trivial as $\sum_{i=1}^{1} f_i=f_1$ which is $\mathcal{A}$-measurable

   $n\in\mathcal{N}\Rightarrow n+1\in\mathcal{N}$. assume that $\{f_i: X\rightarrow\mathbb{R}\}_{i\in\{1,\ldots,n+1\}}$ is $\mathcal{A}$-measurable then $\sum_{i=1}^{n} f_i$ is $\mathcal{A}$-measurable (as $n\in\mathcal{N}$). Hence $\sum_{i=1}^{n+1} f_i=(\sum_{i=1}^{n} f_i)+f_{n+1}$ is $\mathcal{A}$-measurable, proving that $n+1\in\mathcal{N}$.

4. We have for $c\in\mathbb{R}$ then if $c=0$ we have $\forall x\in\mathrm{dom}(f)$ that $(c\cdot f)(x)=0\cdot f(x)=0$ so that $c\cdot f$ is a constant function which is by (1) $\mathcal{A}$-measurable. So we should check only the cases where $c\neq 0$. Then we have two cases to consider:

   $0<c$. then $\forall a\in\mathbb{R}$

$$\begin{aligned} x\in\{c\cdot f<a\} &\Leftrightarrow (c\cdot f)(x)<a\\ &\Leftrightarrow c\cdot f(x)<a\\ &\Leftrightarrow f(x)<\frac{a}{c}\\ &\Leftrightarrow x\in\left\{f<\frac{a}{c}\right\}\end{aligned}$$

   proving that $\{c\cdot f<a\}=\left\{f<\frac{a}{c}\right\}\in\mathcal{A}_{|\mathrm{dom}(f)}$ (as f is $\mathcal{A}$-measurable). Using 18.98 it follows then that $c\cdot f$ is $\mathcal{A}$-measurable.

   $c<0$. then $\forall a\in\mathbb{R}$

$$\begin{aligned} x\in\{c\cdot f<a\} &\Leftrightarrow (c\cdot f)(x)<a\\ &\Leftrightarrow c\cdot f(x)<a\\ &\Leftrightarrow \frac{a}{c}<f(x)\\ &\Leftrightarrow x\in\left\{\frac{a}{c}<f\right\}\end{aligned}$$

   proving that $\{c\cdot f<a\}=\left\{\frac{a}{c}<f\right\}\in\mathcal{A}_{|\mathrm{dom}(f)}$ (as f is $\mathcal{A}$-measurable). Using 18.98 it follows then that $c\cdot f$ is $\mathcal{A}$-measurable.

5. Let $a\in\mathbb{R}$. Define now

$$Q_a=\{(q_1,q_2,q_3,q_4)\in\mathbb{Q}^4|\forall u\in]q_1,q_2[\wedge\forall v\in]q_3,q_4[\text{ we have } u\cdot v<a\}\subseteq\mathbb{Q}^4$$

   then from 9.38,5.66 and 5.101 it follows that

$$\forall a\in\mathbb{R}\text{ we have that } Q_a \text{ is countable} \tag{18.166}$$

   Now given $q\in\mathbb{Q}$ we have as f,g are $\mathcal{A}$-measurable that $\{f<q\},\{q<f\}\in\mathcal{A}_{|\mathrm{dom}(f)}$ and $\{g<q\},\{q<g\}\in\mathcal{A}_{|\mathrm{dom}(g)}$ so that there exists F_q, F'_q, G_q, $G'_q\in\mathcal{A}$ such that

$$\{f<q\}=F_q\bigcap\mathrm{dom}(f)\wedge\{q<f\}=F'_q\bigcap\mathrm{dom}(f)\wedge\{g<q\}=G_q\bigcap\mathrm{dom}(g)\wedge \{q<g\}=G'_q\bigcap\mathrm{dom}(g) \tag{18.167}$$

   Now define

$$\forall(q_1,q_2,q_3,q_4)\in Q_a\quad E_{q_1,q_2,q_3,q_4}=\left\{x\in\mathrm{dom}(f)\bigcap\mathrm{dom}(g)|f(x)\in]q_1,q_2[\wedge g(x)\in]q_3,q_4[\right\} \tag{18.168}$$

then we have

$$
\begin{array}{rcl}
x\in E_{q_1,q_2,q_3,q_4} & \Leftrightarrow & x\in\mathrm{dom}(f)\bigcap\mathrm{dom}(g)\wedge q_1<f(x)\wedge f(x)<q_2\wedge q_3<g(x)\wedge\\
 & & g(x)<q_4\\
 & \Leftrightarrow & x\in\mathrm{dom}(f)\bigcap\mathrm{dom}(g)\wedge x\in\{q_1<f\}\wedge x\in\{f<q_2\}\wedge x\in\\
 & & \{q_3<g\}\wedge(g<q_4)\\
 & \Leftrightarrow & \left(\mathrm{dom}(f)\bigcap\mathrm{dom}(g)\right)\bigcap F'_{q_1}\bigcap F_{q_2}\bigcap G'_{q_3}\bigcap Q_{q_4}
\end{array}
$$

proving that

$$E_{q_1,q_2,q_3,q_4}=\left(\mathrm{dom}(f)\bigcap\mathrm{dom}(g)\right)\bigcap F'_{q_1}\bigcap F_{q_2}\bigcap G'_{q_3}\bigcap Q'_{q_4}\in\mathcal{A}_{|\mathrm{dom}(f)\bigcap\mathrm{dom}(g)} \tag{18.169}$$

As Q_a is countable (see (18.166)) it follows using 18.6 and the above that

$$\bigcup_{(q_1,q_2,q_3,q_4)\in Q_a} E_{q_1,q_2,q_3,q_4}\in\mathcal{A}_{|\mathrm{dom}(f)\bigcap\mathrm{dom}(g)} \tag{18.170}$$

If now $x\in\{f\cdot g<a\}$ then $x\in\mathrm{dom}(f)\bigcap\mathrm{dom}(g)$ and $f(x)\cdot g(x)=(f\cdot g)(x)<a\Rightarrow 0<a-f(x)\cdot g(x)$ so that

$$\eta_x=\min\left(1,\frac{a-f(x)\cdot g(x)}{1+|f(x)|\cdot|g(x)|}\right)>0 \tag{18.171}$$

As $f(x)-\eta_x<f(x)$ by 9.57 there exists a $q_1\in\mathbb{Q}$ such that $f(x)-\eta_x<q_1<f(x)$, as $f(x)<f(x)+\eta_x$ there exists by 9.57 a $q_2\in\mathbb{Q}$ such that $f(x)<q_2<f(x)+\eta_x$, as $g(x)-\eta_x<g(x)$ there exists by 9.57 a $q_3\in\mathbb{Q}$ such that $g(x)-\eta_x<q_3<g(x)$ and finally from $g(x)<g(x)+\eta_x$ there exists a $q_4\in\mathbb{Q}$ such that $g(x)<q_4<g(x)+\eta_q$. To summarize

$$
\begin{array}{ll}
\exists(q_1,q_2,q_3,q_4)\in\mathbb{Q}^4\ \text{ such that} & f(x)-\eta_x<q_1<f(x)\\
 & f(x)<q_2<f(x)+\eta_x\\
 & g(x)-\eta_x<q_3<g(x)\\
 & g(x)<q_4<g(x)+\eta_x
\end{array} \tag{18.172}
$$

Now we have if $r\in]q_1,q_2[$ that $q_1<r<q_2$

$$
\begin{array}{rcl}
r-f(x) & < & q_2-f(x)\\
 & <_{(18.172)} & (f(x)+\eta_x)-f(x)\\
 & = & \eta_x\\
f(x)-r & < & f(x)-q_1\\
 & <_{(18.172)} & f(x)-(f(x)-\eta_x)\\
 & = & \eta_x
\end{array}
$$

proving that

$$\forall r\in]q_1,q[\text{ we have }|r-f(x)|<\eta_x \tag{18.173}$$

If $r \in]q_3, q_4[$ then $q_3 < r < q_4$ so that

$$\begin{array}{rcl} r - g(x) & < & q_4 - g(x) \\ & <_{(18.172)} & (g(x) + \eta_x) - g(x) \\ & = & \eta_x \\ g(x) - r & < & g(x) - q_3 \\ & <_{(18.172)} & g(x) - (g(x) - \eta_x) \\ & = & \eta_x \end{array}$$

proving that

$$\forall r \in]q_3, q_4[\text{ we have } |r - g(x)| < \eta_x \tag{18.174}$$

Next if $r \in]q_1, q_2[$ and $s \in]q_3, q_4[$ then

$$\begin{array}{rcl} r \cdot s - f(x) \cdot g(x) & = & (r - f(x)) \cdot s + f(x) \cdot (s - g(x)) \\ & = & (r - f(x)) \cdot (s - g(x) + g(x)) + f(x) \cdot (s - g(x)) \\ & = & (r - f(x)) \cdot (s - g(x)) + (r - f(x)) \cdot g(x) + f(x) \cdot (s - g(x)) \\ & \leqslant & |r - f(x)| \cdot |s - g(x)| + |r - f(x)| \cdot |g(x)| + |f(x)| \cdot |s - g(x)| \\ & <_{(18.174),(18.173)} & \eta_x^2 + \eta_x \cdot |g(x)| + \eta_x \cdot |f(x)| \\ & = & \eta_x(\eta_x + |g(x)| + |f(x)|) \\ & \leqslant_{(18.171)} & \eta_x(1 + |g(x)| + |f(x)|) \\ & \leqslant_{(18.171)} & \dfrac{a - f(x) \cdot g(x)}{1 + |f(x)| \cdot |g(x)|} \cdot (1 + |g(x)| + |f(x)|) \\ & = & a - f(x) \cdot g(x) \end{array}$$

from which it follows that $r \cdot s < a$ and thus

$$(q_1, q_2, q_3, q_4) \in Q_a \tag{18.175}$$

As from (18.172) $f(x) \in]q_1, q_2[$ and $g(x) \in]q_3, q_4[$ we have using the above and the definition of E_{q_1,q_2,q_3,q_4} that $x \in E_{q_1,q_2,q_3,q_4}$. Which proves that

$$\{f \cdot g < a\} \subseteq \bigcup_{(q_1,q_2,q_3,q_4) \in Q_a} E_{q_1,q_2,q_3,q_4} \tag{18.176}$$

For the opposite inclusion we have

$$\begin{array}{rcl} x \in \bigcup_{(q_1,q_2,q_3,q_4) \in Q_a} E_{q_1,q_2,q_3,q_4} & \Rightarrow & \exists (q_1, q_2, q_3, q_4) \in Q_a \text{ such that } x \in E_{q_1,q_2,q_2,q_4} \\ & \Rightarrow & x \in D \bigcap E \wedge f(x) \in]q_1, q_2[\wedge g(x) \in]q_3, q_4[\\ & \underset{\text{definition of } Q_a}{\Rightarrow} & x \in D \bigcap E \wedge f(x) \cdot g(x) < a \\ & \Rightarrow & x \in \{f \cdot g < a\} \end{array}$$

so using (18.176) we have that $\{f\cdot g<a\}=\bigcup_{(q_1,q_2,q_3,q_4)\in Q_a} E_{q_1,q_2,q_3,q_4}\in\mathcal{A}_{|\mathrm{dom}(f)\bigcap\mathrm{dom}(g)}$ (using (18.170)). As $a\in\mathbb{R}$ was chosen arbitrary we have by 18.98 that

$$f\cdot g\,\text{is}\,\mathcal{A}\text{-measurable}$$

6. Using (4) it is enough to prove that $\frac{1}{g}:\mathrm{dom}(g)\bigcap\{x\in\mathrm{dom}(g)|g(x)\neq 0\}\rightarrow\mathbb{R}$ is $\mathcal{A}$-measurable. Let now $a\in\mathbb{R}$ then we have for a the following cases

$0<a$. then

$$\begin{aligned}
x\in\left\{\frac{1}{g}<a\right\} &\Leftrightarrow \{x\in\mathrm{dom}(g)|g(x)\neq 0\}\text{ and }\frac{1}{g(x)}<a\\
&\Leftrightarrow \{x\in\mathrm{dom}(g)|g(x)\neq 0\}\text{ and }\frac{1}{g(x)}<a\wedge(g(x)<0\vee 0<g(x))\\
&\Leftrightarrow \{x\in\mathrm{dom}(g)|g(x)\neq 0\}\text{ and }\left(\frac{1}{g(x)}<a\wedge g(x)<0\right)\vee\left(\frac{1}{g(x)}<a\wedge 0<g(x)\right)\\
&\Leftrightarrow \{x\in\mathrm{dom}(g)|g(x)\neq 0\}\text{ and }(a\cdot g(x)<1\wedge g(x)<0)\vee(1<a\cdot g(x)\wedge 0<g(x))\\
&\Leftrightarrow \{x\in\mathrm{dom}(g)|g(x)\neq 0\}\text{ and }\left(g(x)<\frac{1}{a}\wedge g(x)<0\right)\vee\left(\frac{1}{a}<g(x)\wedge 0<g(x)\right)\\
&\underset{0<a}{\Leftrightarrow} \{x\in\mathrm{dom}(g)|g(x)\neq 0\}\text{ and }(g(x)<0)\vee\left(\frac{1}{a}<g(x)\right)\\
&\Leftrightarrow \{x\in\mathrm{dom}(g)|g(x)\neq 0\}\bigcap\left(\{g<0\}\bigcup\left(\frac{1}{a}<g\right)\right)
\end{aligned}$$

proving that $\left\{\frac{1}{g}<a\right\}=\{x\in\mathrm{dom}(g)|g(x)\neq 0\}\bigcap\left(\{g<0\}\bigcup\left(\frac{1}{a}<g\right)\right)\in\mathcal{A}_{|\{x\in\mathrm{dom}(g)|g(x)\neq 0\}}$ [see 18.100 and $\mathrm{dom}(f\bigcap\{x\in\mathrm{dom}(g)|g(x)\neq 0\}=\{x\in\mathrm{dom}(g)|g(x)\neq 0\}$]. Hence using 18.98 we have that $\left\{\frac{1}{g}<a\right\}\in\mathcal{A}_{|\mathrm{dom}(g)\bigcap\{x\in\mathrm{dom}(g)|g(x)\neq 0\}}$

$0=a$. then

$$\begin{aligned}
x\in\left\{\frac{1}{g}<a\right\} &\Leftrightarrow x\in\{x\in\mathrm{dom}(g)|g(x)\neq 0\}\text{ and }\frac{1}{g(x)}<0\\
&\Leftrightarrow x\in\{x\in\mathrm{dom}(g)|g(x)\neq 0\}\text{ and }g(x)<0\\
&\Leftrightarrow x\in\{x\in\mathrm{dom}(g)|g(x)\neq 0\}\bigcap\{g<0\}
\end{aligned}$$

proving (see 18.100) that $\left\{\frac{1}{g}<a\right\}=\{x\in\mathrm{dom}(g)|g(x)\neq 0\}\bigcap\{g<0\}\in\mathcal{A}_{|\{x\in\mathrm{dom}(g)|g(x)\neq 0\}}$

$a<0$. then

$$x\in\left\{\frac{1}{g}<a\right\} \Leftrightarrow \{x\in\mathrm{dom}(g)|g(x)\neq 0\} \text{ and } \frac{1}{g(x)}<a$$

$$\underset{\frac{1}{g(x)}<a<0\Rightarrow g(x)<0}{\Leftrightarrow} \{x\in\mathrm{dom}(g)|g(x)\neq 0\} \text{ and } a\cdot g(x)<1\wedge g(x)<0$$

$$\Leftrightarrow \{x\in\mathrm{dom}(g)|g(x)\neq 0\} \text{ and } \frac{1}{a}<g(x)<0$$

$$\Leftrightarrow \{x\in\mathrm{dom}(g)|g(x)\neq 0\}\bigcap\left\{\frac{1}{a}<g\right\}\bigcap\{g<0\}$$

proving (see 18.100) that $\left\{\frac{1}{g}<a\right\}=\{x\in\mathrm{dom}(g)|g(x)\neq 0\}\bigcap\left\{\frac{1}{a}<g\right\}\bigcap\{g<0\}\in\mathcal{A}_{|E\bigcap\{x\in\mathrm{dom}(g)|g(x)\neq 0\}}$

So in all cases we have that $\left\{\frac{1}{g}<a\right\}\in\mathcal{A}_{|\{x\in\mathrm{dom}(g)|g(x)\neq 0\}}$ which as $a\in\mathbb{R}$ is chosen arbitrary by 18.98 proves that

$$\frac{1}{g}\,\text{is Borel measurable}$$

7. Let $a\in\mathbb{R}$

$$\begin{aligned} x\in\{f_{|C}<a\} &\Leftrightarrow x\in\mathrm{dom}(f)\bigcap C\wedge f_{|C}(x)<a\\ &\Leftrightarrow x\in\mathrm{dom}(f)\bigcap C\wedge f(x)<a\\ &\Leftrightarrow x\in\left(\mathrm{dom}(f)\bigcap C\right)\wedge x\in\{f<a\}\\ &\Leftrightarrow x\in\left(\mathrm{dom}(f)\bigcap C\right)\bigcap\{f<a\}\end{aligned}$$

proving that $\{f_{|C}<a\}=C\bigcap\{f<a\}\in\mathcal{A}_{|\mathrm{dom}(f)\bigcap C}\underset{\mathrm{dom}(f)\bigcap C=\mathrm{dom}(f_{|C})}{=}\mathcal{A}_{|\mathrm{dom}(f_{|C})}$

8. As $\forall x\in\mathrm{dom}(f)=\mathrm{dom}(|f|)$ we have $|f|(x)=|f(x)|=(|.|(f(x)))$ it follows that $|f|=|.|\circ f$. Using 18.96 we have as $|.|$ is $\mathcal{B}[\mathbb{R}]$-measurable that $|f|$ is $\mathcal{A}/\mathcal{B}[\mathbb{R}]$-measurable hence $\mathcal{A}$-measurable. □

Theorem 18.104. *Let $\langle X,\mathcal{A}\rangle$ be a measurable space and $\{f_n: X\rightarrow\mathbb{R}\}_{n\in\mathbb{N}}$ a sequence of partial functions such that f_i is $\mathcal{A}$-measurable then we have*

1. $\lim_{n\rightarrow\infty} f_n$ *[see 18.88] is $\mathcal{A}$-measurable. Note that* $D=\mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right)=\left\{x\in\bigcup_{n\in\mathbb{N}}\left(\bigcap_{i\in\{1,\ldots,n\}}\mathrm{dom}(f_i)\right)|\exists k\in\mathbb{N}\text{ such that }\{f_n(x)\}_{n\in\{k,\ldots\infty\}}\text{ has a limit in }\mathbb{R}\right\}$
2. $\sup_{i\in\mathbb{N}} f_i: X\rightarrow\mathbb{R}$ *[see 18.88] is $\mathcal{A}$-measurable. Note that* $D=\mathrm{dom}\left(\sup_{i\in\mathbb{N}} f_i\right)=\{x\in\bigcap_{i\in\mathbb{N}}\mathrm{dom}(f_i)|\sup(\{f_n(x)|n\in\mathbb{N}\})\in\mathbb{R}\}$
3. $\inf_{i\in\mathbb{N}} f_i: X\rightarrow\mathbb{R}$ *[see 18.88] is $\mathcal{A}$-measurable. Note that* $D=\mathrm{dom}\left(\inf_{i\in\mathbb{N}}\right)=\{x\in\bigcap_{i\in\mathbb{N}}\mathrm{dom}(f_i)|\inf(\{f_n(x)|n\in\mathbb{N}\})\in\mathbb{R}\}$

4. $\liminf_{i\rightarrow\infty} f_i: X\rightarrow\mathbb{R}$ *[see 18.88]. Note that* $D=\operatorname{dom}\left(\liminf_{i\rightarrow\infty} f_i\right)=\left\{x\in\bigcap_{i\in\mathbb{N}}\operatorname{dom}(f_i)|\liminf_{i\rightarrow\infty} f_i(x)\in\mathbb{R}\right\}$.
5. $\limsup_{i\rightarrow\infty} f_i: X\rightarrow\mathbb{R}$ *[see 18.88]. Note that* $D=\operatorname{dom}\left(\limsup_{i\rightarrow\infty} f_i\right)=\left\{x\in\bigcap_{i\in\mathbb{N}}\operatorname{dom}(f_i)|\limsup_{i\rightarrow\infty} f_i(x)\in\mathbb{R}\right\}$.

Proof. Let $n\in\mathbb{N}$ and $a\in\mathbb{R}$ then as $f_n: X\rightarrow\mathbb{R}$ is $\mathcal{A}$-measurable, there exists a $F_n(a)\in\mathcal{A}$ such that $\{f_n\leqslant a\}=F_n(a)\bigcap\operatorname{dom}(f_n)$. We proceed now to prove points 1-5.

1. Let $a\in\mathbb{R}$ and take $x\in\left\{\lim_{i\rightarrow\infty} f_i\leqslant a\right\}$ then $x\in D$ and $\left(\lim_{i\rightarrow} f_i\right)(x)\leqslant a$. In other words there exists a $M_x\in\mathbb{N}$ such that $x\in\bigcap_{i\in\{M_x,\ldots,\infty\}}\operatorname{dom}(f_i)$ and $\{f_i(x)\}_{i\in\{M,\ldots,\infty\}}$ has a limit $l=\lim_{i\rightarrow\infty} f_i(x)$ where $l\underset{\text{def}}{=}\left(\lim_{i\rightarrow\infty} f_i\right)(x)\leqslant a$. Take now $k\in\mathbb{N}$ then $l\leqslant a<a+\frac{1}{k}$ and if we take $\varepsilon_k=a+\frac{1}{k}-l>0$ there exists a $n_k\in\{M_x,\ldots,\infty\}$ such that $\forall i\in\{n_k,\ldots,\infty\}$ we have $f_i(x)-l\leqslant|l-f_i(x)|<\varepsilon_k=a+\frac{1}{k}-l$ proving that $f_i(x)<a+\frac{1}{k}$ hence $x\in\left\{f_i<a+\frac{1}{k}\right\}\subseteq F_i\left(a+\frac{1}{k}\right)$. So $x\in\bigcap_{i\in\{n_k,\ldots,\infty\}} F_i\left(a+\frac{1}{k}\right)$ and as $n_k\in\{M_x,\ldots,\infty\}\subseteq\mathbb{N}$ we have $x\in\bigcup_{n\in\mathbb{N}}\left(\bigcap_{i\in\{n,\ldots,\infty\}} F_i\left(a+\frac{1}{k}\right)\right)$. Further as k was chosen arbitrary and $x\in D$ we have that

$$\left\{\lim_{i\rightarrow\infty} f_i\leqslant a\right\}\subseteq D\bigcap\left(\bigcap_{k\in\mathbb{N}}\left(\bigcup_{n\in\mathbb{N}}\left(\bigcap_{i\in\{n,\ldots,\infty\}} F_i\left(a+\frac{1}{k}\right)\right)\right)\right)\tag{18.177}$$

Let $x\in D\bigcap\left(\bigcap_{k\in\mathbb{N}}\left(\bigcup_{n\in\mathbb{N}}\left(\bigcap_{i\in\{n,\ldots,\infty\}} F_i\left(a+\frac{1}{k}\right)\right)\right)\right)$ then as $x\in D$ there exists a $m\in\mathbb{N}$ such that $x\in\bigcap_{i\in\{m,\ldots,\infty\}}\operatorname{dom}(f_i)$ and for $\{f_i(x)\}_{i\in\{m,\ldots,\infty\}}$ $\lim_{i\rightarrow\infty} f_i(x)\in\mathbb{R}$. As $x\in\bigcap_{k\in\mathbb{N}}\left(\bigcup_{n\in\mathbb{N}}\left(\bigcap_{i\in\{n,\ldots,\infty\}} F_i\left(a+\frac{1}{n}\right)\right)\right)$ we have $\forall k\in\mathbb{N}$ that there exists a $n\in\mathbb{N}$ such that $\forall i\in\{n,\ldots,\infty\}$ we have $x\in F_i\left(a+\frac{1}{k}\right)$. Take now $p=\max(n,m)$ then $\forall k\in\mathbb{N}$ we have for $i\in\{p,\ldots,\infty\}$ that $x\in\operatorname{dom}(f_i)\bigcap F_i\left(a+\frac{1}{k}\right)=\left\{f_i<a+\frac{1}{k}\right\}\Rightarrow f_i(x)<a+\frac{1}{k}$. Using 9.56 we have then that $\forall i\in\{p,\ldots,\infty\}$ $f_i(x)\leqslant a$. So using 12.337 we have that $\left(\lim_{i\rightarrow\infty} f_i\right)(x)=\lim_{i\rightarrow\infty} f_i(x)\leqslant a$ or $x\in\left\{\lim_{i\rightarrow\infty} f_i\leqslant a\right\}$ giving

$$D\bigcap\left(\bigcap_{k\in\mathbb{N}}\left(\bigcup_{n\in\mathbb{N}}\left(\bigcap_{i\in\{n,\ldots,\infty\}} F_i\left(a+\frac{1}{k}\right)\right)\right)\right)\subseteq\left\{\lim_{i\rightarrow\infty} f_i\leqslant a\right\}\tag{18.178}$$

which together with (18.177) gives

$$D\bigcap\left(\bigcap_{k\in\mathbb{N}}\left(\bigcup_{n\in\mathbb{N}}\left(\bigcap_{i\in\{n,\ldots,\infty\}} F_i\left(a+\frac{1}{k}\right)\right)\right)\right)=\left\{\lim_{i\rightarrow\infty} f_i\leqslant a\right\}\tag{18.179}$$

Using 1.107 we have then

$$\bigcap_{k\in\mathbb{N}}\left(\bigcup_{n\in\mathbb{N}}\left(\bigcap_{i\in\{n,\ldots,\infty\}}\left(D\bigcap F_i\left(a+\frac{1}{k}\right)\right)\right)\right)=\left\{\lim_{i\rightarrow\infty} f_i\leqslant a\right\}\tag{18.180}$$

Now $D\bigcap F_i\big(a+\frac{1}{k}\big)\in\mathcal{A}_{|D}$, hence using 18.8 we have $\bigcap_{i\in\{n,\ldots,\infty\}}\big(D\bigcap F_i\big(a+\frac{1}{k}\big)\big)\in\mathcal{A}_{|D}$, so $\bigcup_{n\in\mathbb{N}}\big(\bigcap_{i\in\{n,\ldots,\infty\}}\big(D\bigcap F_i\big(a+\frac{1}{k}\big)\big)\big)\in\mathcal{A}_{|D}$ and thus $\bigcap_{k\in\mathbb{N}}\big(\bigcup_{n\in\mathbb{N}}\big(\bigcap_{i\in\{n,\ldots,\infty\}}\big(D\bigcap F_i\big(a+\frac{1}{k}\big)\big)\big)\big)\in\mathcal{A}_{|D}$ which proves that $\left\{\lim_{i\rightarrow\infty}f_i\leqslant a\right\}\in\mathcal{A}_{|D}$. So by 18.98 we have that

$$\lim_{i\rightarrow\infty}f_i\text{ is }\mathcal{A}\text{-measurable}$$

2. Let $x\in\left\{\sup_{i\in\mathbb{N}}f_i\leqslant a\right\}$ then $x\in D$ so that $x\in\bigcap_{i\in\mathbb{N}}\operatorname{dom}(f_i)$, $\left(\sup_{i\in\mathbb{N}}f_i\right)(x)=\sup\left(\{f_i(x)|i\in\mathbb{N}\}\right)\in\mathbb{R}$ and $\sup(\{f_i(x)|i\in\mathbb{N}\})\leqslant a$. Hence $\forall i\in\mathbb{N}$ we have $f_i(x)\leqslant a\Rightarrow x\in\{f_i\leqslant a\}\subseteq F_i(a)$ proving

$$\left\{\sup_{i\in\mathbb{N}}f_i\leqslant a\right\}\subseteq D\bigcap\left(\bigcap_{i\in\mathbb{N}}F_i(a)\right)\tag{18.181}$$

If $x\in D\bigcap\left(\bigcap_{i\in\mathbb{N}}F_i(a)\right)$ then $x\in D\Rightarrow x\in\bigcap_{i\in\mathbb{N}}\operatorname{dom}(f_i)$ and $x\in\bigcap_{i\in\mathbb{N}}F_i(a)$ so that $\forall i\in\mathbb{N}$ we have $x\in F_i(a)\bigcap\operatorname{dom}(f_i)=\{f_i\leqslant a\}\Rightarrow f_i(x)\leqslant a$. Hence $\left(\sup_{i\in\mathbb{N}}f_i\right)(x)=\sup\left(\{f_i(x)|i\in\mathbb{N}\}\right)\leqslant a$ or $x\in\left\{\sup_{i\in\mathbb{N}}f_i\leqslant a\right\}$ proving that $D\bigcap\left(\bigcap_{i\in\mathbb{N}}F_i(a)\right)\subseteq\left\{\sup_{i\in\mathbb{N}}f_i\leqslant a\right\}$. Using this in (18.181)

$$\left\{\sup_{i\in\mathbb{N}}f_i\leqslant a\right\}=D\bigcap\left(\bigcap_{i\in\mathbb{N}}F_i(a)\right)\tag{18.182}$$

As $D\bigcap\left(\bigcap_{i\in\mathbb{N}}F_i(a)\right)=\bigcap_{i\in\mathbb{N}}\left(D\bigcap F_i(A)\right)\in\mathcal{A}_{|D}$ we have $\left\{\sup_{i\in\mathbb{N}}f_i\leqslant a\right\}\in\mathcal{A}_{|D}$. Applying then 18.98 give that

$$\sup_{i\in\mathbb{N}}f_i\text{ is }\mathcal{A}\text{-measurable}$$

3. As $\forall i\in\mathbb{N}$ f_i is $\mathcal{A}$-measurable we have by 18.103 (3) that $-f_i$ is $\mathcal{A}$-measurable. So by (2) $\sup_{i\in\mathbb{N}}(-f_i)$ is $\mathcal{A}$-measurable and using 18.103 (3) again we have that $-\sup_{i\in\mathbb{N}}(-f_i)$ is $\mathcal{A}$-measurable. Further as

$$\begin{aligned}\left(\sup_{i\in\mathbb{N}}(-f_i)\right)(x) &= \sup\left(\{-f_i(x)|i\in\mathbb{N}\}\right)\\ &\underset{\text{Algorithm 17.30}}{=} -\inf\left(\{f_i(x)|i\in\mathbb{N}\}\right)\\ &= -\left(\inf_{i\in\mathbb{N}}f_i\right)(x)\end{aligned}$$

which proves that $\inf_{i\in\mathbb{N}}f_i=-\sup_{i\in\mathbb{N}}f_i$. Hence

$$\inf_{i\in\mathbb{N}}f_i\text{ is }\mathcal{A}\text{-measurable}$$

4. Let $x\in\bigcap_{i\in\mathbb{N}}\mathrm{dom}(f_i)$ then

$$\begin{aligned}\left(\limsup_{i\to\infty} f_i\right)(x) &= \limsup_{i\to\infty} f_i(x)\\ &= \inf\left(\{\sup\left(\{f_i(x)|i\in\{n,\ldots,\infty\}\}\right)|n\in\mathbb{N}\}\right)\\ &= \inf\left(\{\sup\left(\{f_{i+n-1}(x)|i\in\mathbb{N}\}\right)|n\in\mathbb{N}\}\right)\\ &= \inf\left(\left(\sup_{i\in\mathbb{N}} f_{i+n-1}\right)(x)|n\in\mathbb{N}\right)\\ &= \left(\inf_{n\in\mathbb{N}}\left(\sup_{i\in\mathbb{N}} f_{i+n-1}\right)\right)(x)\end{aligned}$$

proving that $\limsup_{i\to\infty} f_i=\inf_{n\in\mathbb{N}}\left(\sup_{i\in\mathbb{N}} f_{i+n-1}\right)$. Let $n\in\mathbb{N}$ then as $\forall i\in\mathbb{N}$ $i+n-1\in\mathbb{N}$ we have that f_{i+n-1} is $\mathcal{A}$-measurable, so that by (2) $\sup_{i\in\mathbb{N}} f_{i+n-1}$ is $\mathcal{A}$-measurable. Finally by (3) $\inf_{n\in\mathbb{N}}\left(\sup_{i\in\mathbb{N}} f_{i+n-1}\right)$ is $\mathcal{A}$-measurable proving that

$$\limsup_{i\to\infty} f_i \text{ is } \mathcal{A}\text{-measurable}$$

5. Let $x\in\bigcap_{i\in\mathbb{N}}\mathrm{dom}(f_i)$ then

$$\begin{aligned}\left(\liminf_{i\to\infty} f_i\right)(x) &= \liminf_{i\to\infty} f_i(x)\\ &= \sup\left(\{\inf\left(\{f_i(x)|i\in\{n,\ldots,\infty\}\}\right)|n\in\mathbb{N}\}\right)\\ &= \sup\left(\{\inf\left(\{f_{i+n-1}(x)|i\in\mathbb{N}\}\right)|n\in\mathbb{N}\}\right)\\ &= \sup\left(\left(\inf_{i\in\mathbb{N}} f_{i+n-1}\right)(x)|n\in\mathbb{N}\right)\\ &= \left(\sup_{n\in\mathbb{N}}\left(\inf_{i\in\mathbb{N}} f_{i+n-1}\right)\right)(x)\end{aligned}$$

proving that $\liminf_{i\to\infty} f_i=\sup_{n\in\mathbb{N}}\left(\inf_{i\in\mathbb{N}} f_{i+n-1}\right)$. Let $n\in\mathbb{N}$ then as $\forall i\in\mathbb{N}$ $i+n-1\in\mathbb{N}$ we have that f_{i+n-1} is $\mathcal{A}$-measurable, so that by (3) $\inf_{i\in\mathbb{N}} f_{i+n-1}$ is $\mathcal{A}$-measurable. Finally by (2) $\sup_{n\in\mathbb{N}}\left(\inf_{i\in\mathbb{N}} f_{i+n-1}\right)$ is $\mathcal{A}$-measurable proving that

$$\liminf_{i\to\infty} f_i \text{ is } \mathcal{A}\text{-measurable}$$ □

Proposition 18.105. *Let $\langle X,\mathcal{A}\rangle$ be a measurable space and $f: X\to\mathbb{R}$ a partial function then f is $\mathcal{A}$-measurable if and only if $\exists h: X\to\mathbb{R}$ $\mathcal{A}$-measurable such that $h_{|\mathrm{dom}(f)}=f$ (note here h is a function so $\mathrm{dom}(h)=X$)*

Proof. If $h: X\to\mathbb{R}$ is a $\mathcal{A}$-measurable function such that $h_{|\mathrm{dom}(f)}=f$ then by 18.103 (6) $h_{|\mathrm{dom}(f)}$ is $\mathcal{A}$-measurable hence f is $\mathcal{A}$-measurable. So we must only prove the opposite implication. Let $f: X\to\mathbb{R}$ be a $\mathcal{A}$-measurable partial function. Take $q\in\mathbb{Q}$ then $\{f\leqslant q\}\in\mathcal{A}_{|\mathrm{dom}(f)}$ and thus there exists a $A_q\in\mathcal{A}$ such that

$$\{f\leqslant q\}=A_q\bigcap\mathrm{dom}(f) \tag{18.183}$$

As $\mathbb{Q}$ is denumerable and thus countable (see 9.38) we have by 18.8 that $\bigcup_{q\in Q} A_q \in \mathcal{A}$, hence we have that

$$F = X \setminus \bigcup_{q\in Q} A_q \in \mathcal{A} \tag{18.184}$$

Further if $n\in\mathbb{N}$ we have that $\{q\in\mathbb{Q}|q\leqslant -n\}\subseteq\mathbb{Q}$ is countable (see 5.66) it follows using 18.8 that $\bigcup_{q\in\{q\in\mathbb{Q}|q\leqslant -n\}} A_q\in\mathcal{A}$, hence we have that

$$G = \bigcap_{n\in\mathbb{N}} \left(\bigcup_{q\in\{q\in\mathbb{Q}|q\leqslant -n\}} A_q\right)\in\mathcal{A} \tag{18.185}$$

If $x\in\mathrm{dom}(f)$ then $f(x)<f(x)+1$, using 9.57 there exists a $q\in\mathbb{Q}$ such that $f(x)<q<f(x)+1$, so that $x\in\{f\leqslant q\}\subseteq A_q\subseteq\bigcup_{q\in\mathbb{Q}} A_q$ hence $x\notin F$ proving that $F\bigcap\mathrm{dom}(f)=\emptyset$. Also using 9.55 there exists a $n\in\mathbb{N}$ such that $-f(x)<n\Rightarrow -n<f(x)$ so that $\forall q'\in Q$ with $q'\leqslant -n$ we have $q'<f(x)$. Assume that $x\in A_{q'} \underset{x\in\mathrm{dom}(f)}{\Rightarrow} x\in A_{q'}\bigcap\mathrm{dom}(f)=\{f\leqslant q'\}\Rightarrow f(x)\leqslant q'$ contradicting $q'<f(x)$, hence we must have that $x\notin A_{q'}$. This proves that $x\notin\bigcup_{q\in\{q\in\mathbb{Q}|q\leqslant -n\}} A_q$ and thus $x\notin G$ or $\mathrm{dom}(f)\bigcap G=\emptyset$. To summarize we have

$$\mathrm{dom}(f)\bigcap F=\emptyset \text{ and } \mathrm{dom}(f)\bigcap G=\emptyset \tag{18.186}$$

Set $H=X\setminus(F\bigcup G)$ then using (18.184) and (18.185) we have that $H\in\mathcal{A}$. Now given $x\in H$ we have $x\notin F\Rightarrow\exists q\in\mathbb{Q}$ such that $x\in A_q$ proving that $\{q\in\mathbb{Q}|x\in A_q\}\neq\emptyset$, Further as $x\notin G$ there exists a $n\in\mathbb{N}$ such that $\forall q\in\mathbb{Q}$ with $q\leqslant -n$ we have $x\notin A_q$. So for all $q\in\{q\in\mathbb{Q}|x\in A_q\}$ we must have that $-n<q$ or $\{q\in\mathbb{Q}|x\in\mathcal{A}_q\}$ is bounded below. As $\mathbb{R}$ is conditionally complete (see 9.43) we have

$$\forall x\in H=X\setminus\left(F\bigcup G\right) \quad \inf(\{q\in Q|x\in A_q\})\,\text{exist in}\,\mathbb{R} \tag{18.187}$$

We define now $h: X\rightarrow\mathbb{R}$ by

$$h(x)=\begin{cases}\inf(\{q\in\mathbb{Q}|x\in A_q\})\text{ if } x\in H\\ 0 \text{ if } x\notin H\end{cases} \tag{18.188}$$

Next we prove that h extends f. If $x\in\mathrm{dom}(f)$ then using (18.186) we have that $x\in H$. As for $f(x)$ we can find (see 9.57) a $q\in\mathbb{Q}$ such that $f(x)<q<f(x+1)$ hence $q\in\{f\leqslant q\}\subseteq A_q\Rightarrow q\in\{q\in\mathbb{Q}|x\in A_q\}$ so that $f(x)\leqslant\inf(\{q\in\mathbb{Q}|x\in A_q\})=h(x)$ proving that

$$f(x)\leqslant h(x). \tag{18.189}$$

If $\varepsilon>0$ then using 9.57 there a $q\in\mathbb{Q}$ such that $f(x)<q<f(x)+\varepsilon$, as $f(x)<q$ gives $x\in\{f\leqslant q\}\subseteq A_q$, it follows from the definition of infinum that $h(x)\leqslant q<f(x)+\varepsilon$, proving that $h(x)<f(x)+\varepsilon$. As $\varepsilon>0$ was chosen arbitrary it follows from 9.56 that $h(x)\leqslant f(x)$ giving using (18.189) that $f(x)=h(x)$. Hence

$$\forall x\in\mathrm{dom}(f)\,\text{we have}\,h(x)=f(x) \text{ or } h_{|\mathrm{dom}(f)}=f \tag{18.190}$$

Finally we must prove that h is measurable. Let $a \in \mathbb{R}$ then we must consider the following cases:

0<a. Let $x \in \{x \in X | h(x) < a\}$ and consider for x the following cases

$x \in F \bigcup G$. then $x \in (F \bigcup G) \subseteq \left(H \bigcap \bigcup_{q \in \{q \in \mathbb{Q} | q < a\}} A_q\right) \bigcup (F \bigcup G)$

$x \notin F \bigcup G$. then $x \in H$ hence $h(x) = \inf(\{q \in \mathbb{Q} | x \in A_q\}) < a$, so there exists a $q \in \mathbb{Q} \vdash x \in A_q$ such that $h(x) \leqslant q < a$, giving $x \in \bigcup_{q \in \{q \in \mathbb{Q} | q < a\}} A_q$, so $x \in H \bigcap \bigcup_{q \in \{q \in \mathbb{Q} | q < a\}} A_q \subseteq \left(H \bigcap \bigcup_{q \in \{q \in \mathbb{Q} | q < a\}} A_q\right) \bigcup (F \bigcup G)$

hence

$$\{x \in X | h(x) < a\} \subseteq \left(H \bigcap \bigcup_{q \in \{q \in \mathbb{Q} | q < a\}} A_q\right) \bigcup (F \bigcup G) \tag{18.191}$$

If $x \in \left(H \bigcap \bigcup_{q \in \{q \in \mathbb{Q} | q < a\}} A_q\right) \bigcup (F \bigcup G)$ then either

$x \in F \bigcup G$. then $x \notin H$ so that $h(x) = 0 < a$ proving that $x \in \{x \in X | h(x) < a\}$

$x \notin F \bigcup G$. then $x \in H \bigcap \bigcup_{q \in \{q \in \mathbb{Q} | q < a\}} A_q$ hence $x \in H$ and $\exists q \in Q$ with $q < a$ and $x \in A_q$. From $x \in H$ we have that $h(x) = \inf(\{q' \in Q | x \in A_{q'}\})$ so that $h(x) \leqslant q < a$ which proves that $x \in \{x \in X | h(x) < a\}$

So in all cases $x \in \{x \in X | h(x) < a\}$ and thus we have $\left(H \bigcap \bigcup_{q \in \{q \in \mathbb{Q} | q < a\}} A_q\right) \bigcup (F \bigcup G) \subseteq \{x \in X | h(x) < a\}$ which together with (18.191) gives

$$\{x \in X | h(x) < a\} = \left(H \bigcap \bigcup_{q \in \{q \in \mathbb{Q} | q < a\}} A_q\right) \bigcup (F \bigcup G) \tag{18.192}$$

As $\{a \in \mathbb{Q} | q < a\} \subseteq \mathbb{Q}$ is countable [see 9.38, 5.66] and $A_q \in \mathcal{A}$ we have by 18.8 that $\bigcup_{q \in \{q \in \mathbb{Q} | q < a\}} A_q \in \mathcal{A}$. Further $F, G \in \mathcal{A}$ [see (18.184),(18.185)] and thus also $H = X \setminus (F \bigcup G) \in \mathcal{A}$. To summarize

$$\bigcup_{q \in \{q \in \mathbb{Q} | q < a\}} A_q, F, G, H \in \mathcal{A} \tag{18.193}$$

So using (18.192) we have

$$\{x \in X | h(x) < a\} \in \mathcal{A} \tag{18.194}$$

$a \leqslant 0$. If $x \in \{x \in X | h(x) < a\}$ then $h(x) < a$. If $x \notin H$ then $0 = h(x) < a \leqslant 0$ giving the contradiction $0 < 0$. So we must have that $x \in H$, then $h(x) = \inf(\{q \in \mathbb{Q} | x \in A_q\})$ and thus there exists a $q \in \mathbb{Q}$ with $x \in A_q$ such that $h(x) \leqslant q < a$. So $x \in \bigcup_{q \in \{q \in Q | q < a\}} A_q$ giving

$$\{x \in X | h(x) < a\} \subseteq H \bigcap \left(\bigcup_{q \in \{q \in \mathbb{Q} | q < a\}} A_q\right) \tag{18.195}$$

If $x\in H\bigcap\left(\bigcup_{q\in\{q\in\mathbb{Q}|q<a\}}A_q\right)$ then $x\in X$ and $\exists q\in\mathbb{Q}$ with $q<a$ such that $x\in A_q$. As $x\in H$ we have that $h(x)=\inf(\{q'\in\mathbb{Q}|x\in A_{q'}\})\leqslant q<a$ proving that $x\in\{x\in X|h(x)<a\}$. So $H\bigcap\left(\bigcup_{q\in\{q\in\mathbb{Q}|q<a\}}A_q\right)\subseteq\{x\in X|h(x)<a\}$ which by (18.195) proves

$$\{x\in X|h(x)<a\}=H\bigcap\left(\bigcup_{q\in\{q\in\mathbb{Q}|q<a\}}A_q\right)\tag{18.196}$$

Using (18.193) on the above gives

$$\{x\in X|h(x)<a\}$$

So we have proves that $\forall a\in\mathbb{R}$ we have $\{h<a\}=\{x\in X|h(x)<a\}\in\mathcal{A}$ which proves (see 18.98) that

$$h\text{ is }\mathcal{A}\text{-measurable}\tag{18.197}$$

The proposition is then proved by (18.190) and (18.197) □

Theorem 18.106. *Let $n\in\mathbb{N}$, $\langle X,\mathcal{A}\rangle$ a measurable space and $\forall i\in\{1,\ldots,n\}$ $D_i\subseteq X$ and $f_i\colon D_i\rightarrow\mathbb{R}$ $\mathcal{A}$-measurable partial functions. Then if we take $f\colon X\rightarrow\mathbb{R}^n$ defined by $f(x)=(f_1(x),\ldots,f_n(x))$ $\forall x\in\bigcap_{i\in\{1,\ldots,n\}}\operatorname{dom}(f_i)$ [or $\operatorname{dom}(f)=\bigcap_{i\in\{1,\ldots,n\}}\operatorname{dom}(f_i)$] we have*

1. *$\forall B\in\mathcal{B}[\mathbb{R}^n]$ $f^{-1}(B)\in\mathcal{A}_{|\bigcap_{i\in\{1,\ldots,n\}}\operatorname{dom}(f_i)}$ [or f is $\mathcal{A}/\mathcal{B}[\mathbb{R}^n]$-measurable.*
2. *If $A\subseteq\mathbb{R}^n$ and $h\colon A\rightarrow\mathbb{R}$ a $\mathcal{A}$-measurable function then $h\circ f\colon D\bigcap h^{-1}(A)\rightarrow\mathbb{R}$ is Borel measurable.*

Proof.

1. Define $D=\bigcap_{i\in\{1,\ldots,n\}}\operatorname{dom}(f_i)$

$$\mathcal{C}=\{B\in\mathbb{R}^n|f^{-1}(B)\in\mathcal{A}_{|D}\}\tag{18.198}$$

As $f^{-1}(\emptyset)=\emptyset\in\mathcal{A}_{|D}$ it follows that

$$\emptyset\in\mathcal{C}\tag{18.199}$$

If $B\in\mathcal{C}$ then $f^{-1}(B)\in\mathcal{A}_{|D}$ a σ-algebra on D then $f^{-1}(\mathbb{R}^n\setminus B)\underset{2.54(4)}{=}D\setminus f^{-1}(B)\in\mathcal{A}_{|D}$ [see 18.8] proving

$$\forall B\in\mathcal{C}\text{ we have }D\setminus B\in\mathcal{C}\tag{18.200}$$

Further if $\{B_i\}_{i\in\mathbb{N}}\subseteq\mathcal{C}$ then $\forall i\in\mathbb{N}$ we have $f^{-1}(B_i)\in\mathcal{A}_{|D}$. As $f^{-1}(\bigcup_{i\in\mathbb{N}}B_i)\underset{2.58}{=}\bigcup_{i\in\mathbb{N}}f^{-1}(B_i)\in\mathcal{A}_{|D|}$ we have together with (18.199) and (18.200) that

$$\mathcal{C}\text{ is a }\sigma\text{-algebra}\tag{18.201}$$

Let $H\in\mathcal{H}^n$ there exists a $a\in\mathbb{R}$ and $i\in\{1,\ldots,n\}$ so that $H=]-\infty,a]_i$ then

$$\begin{aligned}x\in f^{-1}(H)&\Leftrightarrow x\in D\wedge f(x)\in]-\infty,a]_i\\&\Leftrightarrow x\in D\wedge f(x)_i\leqslant a\\&\Leftrightarrow x\in D\wedge f_i(x)\leqslant a\\&\Leftrightarrow x\in D\wedge f_i(x)\in]-\infty,a]\\&\Leftrightarrow x\in D\wedge x\in f^{-1}(]-\infty,a])\end{aligned}$$

proving that $f^{-1}(H) = D \bigcap f^{-1}(]-\infty, a])$. As by 18.62 $]-\infty, a] \in \mathcal{B}[\mathbb{R}]$ and f_i is measurable we have that $f^{-1}(]-\infty, a]) \in \mathcal{A}_{|D_i}$ so there exists a $A \in \mathcal{A}$ such that $f^{-1}(]-\infty, a]) = A \bigcap D_i$, hence $f^{-1}(H) = D \bigcap (A \bigcap D_i) \underset{D \subseteq D_i}{=} A \bigcap D$, proving that $f^{-1}(H) \in \mathcal{A}_{|D}$. So we have proved that $\forall H \in \mathcal{H}^n$ we have $f^{-1}(H) \in \mathcal{A}_{|D}$ or that $\mathcal{H}^n \subseteq \mathcal{C}$ hence we have that

$$\sigma(\mathcal{H}^n) \subseteq \mathcal{C} \tag{18.202}$$

As $\sigma[\mathcal{H}^n] \underset{18.86}{=} \mathcal{B}[\mathbb{R}^n]$ we have that $\mathcal{B}[\mathbb{R}^n] \subseteq \mathcal{C}$ so

$$\forall B \in \mathcal{B}[\mathbb{R}^n] \text{ we have } f^{-1}(B) \in \mathcal{A}_{|D} = \mathcal{A}_{|\bigcap_{i \in \{1,\ldots,n\}} \mathrm{dom}(f_i)} \tag{18.203}$$

2. As by (1) f is $\mathcal{A}/\mathcal{B}[\mathbb{R}^n]$-measurable and h is Borel measurable [or $\mathcal{B}[\mathbb{R}^n]/\mathcal{B}[\mathbb{R}]$-measurable] we have by 18.96 that $h \circ f$ is Borel measurable. □

18.2.2 Measure Integral

18.2.2.1 Integral of simple functions

Recap the definition of a characteristic function (see 2.14)

Definition 18.107. *Let X be a set, $A \in \mathcal{P}(X)$ then the* ***characteristics function*** *$\mathcal{X}_A: X \to \mathbb{R}$ is defined by $\mathcal{X}_A(x) = \begin{cases} 1 \text{ if } x \in A \\ 0 \text{ if } x \in X \setminus A \end{cases}$*

Note 18.108. The characteristics function is a function so $\mathrm{dom}(\mathcal{X}_{|A}) = X$

Lemma 18.109. *Let $\langle X, \mathcal{A} \rangle$ be a measurable space then $\forall A \in \mathcal{A}$ we have that $\mathcal{X}_A: X \to \mathbb{R}$ is Borel measurable.*

Proof. Let $a \in \mathbb{R}$ consider then the following cases for

$a = 1$. then $\{\mathcal{X}_A < a\} = A \in \mathcal{A}$

$a < 1$. then $\{\mathcal{X}_A < a\} = X \setminus A \in \mathcal{A}$

$1 < a$. then $\{\mathcal{X}_A < a\} = \emptyset$ □

Lemma 18.110. *Let X be a set then we have the following properties for characteristics functions*

1. *$\forall A, B \in \mathcal{P}(X)$ we have $\mathcal{X}_{A \bigcap B} = \mathcal{X}_A \cdot \mathcal{X}_B$*
2. *Let $n \in \mathbb{N}$ then $\forall \{A_i\}_{i \in} \subseteq \mathcal{P}(X)$ pairwise disjoint we have $\mathcal{X}_{\bigsqcup_{i \in \{1,\ldots,n\}} A_i} = \sum_{i=1}^n \mathcal{X}_{A_i}$*

Proof.

1. Let $x \in X$ then we have the following possibilities for x

   $x \in A \wedge x \in B$. then $(\mathcal{X}_A \cdot \mathcal{X}_B)(x) = \mathcal{X}_A(x) \cdot \mathcal{X}_B(x) = 1 \cdot 1 = 1 \underset{x \in A \bigcap B}{=} \mathcal{X}_{A \bigcap B}(x)$

   $x \in A \wedge x \notin B$. then $(\mathcal{X}_A \cdot \mathcal{X}_B)(x) = \mathcal{X}_A(x) \cdot \mathcal{X}_B(x) = 1 \cdot 0 = 0 \underset{x \notin A \bigcap B}{=} \mathcal{X}_{A \bigcap B}(x)$

$x \notin A \wedge x \in B$. then $(\mathcal{X}_A \cdot \mathcal{X}_B)(x) = \mathcal{X}_A(x) \cdot \mathcal{X}_B(x) = 0 \cdot 1 = 0 \underset{x \notin A \bigcap B}{=} \mathcal{X}_{A \bigcap B}(x)$

$x \notin A \wedge x \notin B$. then $(\mathcal{X}_A \cdot \mathcal{X}_B)(x) = \mathcal{X}_A(x) \cdot \mathcal{X}_B(x) = 0 \cdot 0 = 0 \underset{x \notin A \bigcap B}{=} \mathcal{X}_{A \bigcap B}(x)$

proving that $\mathcal{X}_A \cdot \mathcal{X}_B = \mathcal{X}_{A \bigcap B}$

2. Let $x \in X$ then we have the following possibilities for x

   $x \in \bigsqcup_{i \in \{1,\ldots,n\}} A_i$. then there exists a $i \in \{1, \ldots, n\}$ such that $x \in A_i$ and $\forall j \in \{1, \ldots, n\} \setminus \{i\}$ that $x \in A_j = 0$. Hence $\sum_{j=1}^{n} \mathcal{X}_{A_j}(x) = \mathcal{X}_{A_i}(x) = 1 \underset{x \in \bigsqcup_{i \in \{1,\ldots,n\}} A_i}{=} \mathcal{X}_{\bigsqcup_{i \in \{1,\ldots,n\}} A_i}$

   $x \notin \bigsqcup_{i \in \{1,\ldots,n\}} A_i$. then $\forall i \in \{1, \ldots, n\}$ we have $x \notin A_i \Rightarrow \mathcal{X}_{A_i}(x) = 0$ so that $\sum_{j=1}^{n} \mathcal{X}_{A_j}(x) = 0 \underset{x \notin \bigsqcup_{i \in \{1,\ldots,n\}} A_i}{=} \mathcal{X}_{\bigsqcup_{i \in \{1,\ldots,n\}} A_i}$

   which proves that $\mathcal{X}_{\bigsqcup_{i \in \{1,\ldots,n\}} A_i} = \sum_{i=1}^{n} \mathcal{X}_{A_i}$ □

We define now the concept of a simple function that will be the basis to define the measure integral. We require that a simple function is finite and measurable. This lead to the following definition.

Definition 18.111. *Let $\langle X, \mathcal{A}, \mu \rangle$ be a measure space then a function of the form $\sum_{i=1}^{n} \alpha_i \cdot \mathcal{X}_{A_i}: X \to \mathbb{R}$ where $n \in \mathbb{N}$, $\{\alpha_i\}_{i \in \{1,\ldots,n\}}$, $\{A_i\}_{i \in \{1,\ldots,n\}} \subseteq \mathcal{A}$ such that $\mu(A_i) < \infty$ is called a* ***simple function****. The set of all the simple functions is noted as $\mathcal{S}[X, \mathcal{A}]$.*

Note 18.112. A simple function is a **function** so if $f \in \mathcal{S}[X, \mathcal{A}]$ then $\operatorname{dom}(f) = X$

Definition 18.113. *Let $f \in \mathcal{S}[X, \mathcal{A}]$ then a representation of f consists of $\{A_i\}_{i \in \{1,\ldots,n\}} \subseteq \mathcal{A}$, $\{\alpha_i\}_{i \in \{1,\ldots,n\}} \subseteq \mathbb{R}$, $n \in \mathbb{N}$ such that*

1. $\forall i \in \{1, \ldots, n\}$ $\mu(A_i) < \infty$
2. $f = \sum_{i=1}^{n} \alpha_i \cdot \mathcal{X}_{A_i}$

Note 18.114. Using the definition of a simple function it is clear that every simple function has always a representation. Of course a simple function can have many different representations.

Simple functions have the following properties

Lemma 18.115. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space then we have*

1. *$\forall f \in \mathcal{S}[X, \mathcal{A}]$ we have that f is $\mathcal{A}$-measurable*
2. *$\forall f, g \in \mathcal{S}[X, \mathcal{A}]$ we have that $f + g \in \mathcal{S}[X, \mathcal{A}]$*
3. *If $f \in \mathcal{S}[X, \mathcal{A}]$ and $c \in \mathbb{R}$ then $c \cdot f \in \mathcal{S}[X, \mathcal{A}]$*

Proof.

1. *From the definition of a simple function, the fact that characteristics function are μ-measurable we have using 18.103 that that f is $\mathcal{A}$-measurable.*
2. *Let $f = \sum_{i=1}^{n} \alpha_i \cdot \mathcal{X}_{A_i}$, $g = \sum_{i=1}^{m} \beta_i \cdot \mathcal{X}_{B_i}$ where $\{\alpha_i\}_{i\in\{1,\ldots,n\}}, \{\beta_i\}_{i\in\{1,\ldots,m\}} \subseteq \mathbb{R}$, $\{A_i\}_{i\in\{1,\ldots,n\}}$, $\{B_i\}_{i\in\{1,\ldots,m\}} \subseteq \mathcal{A}$ and $\mu(A_i) < \infty$, $\mu(B_i) < \infty$ then $\sum_{i=1}^{n+m} \varsigma_i \cdot \mathcal{X}_{C_i}$ where $C_i = \begin{cases} A_i \text{ if } i \in \{1, \ldots, n\} \\ B_i \text{ if } i \in \{n+1, \ldots, n+m\} \end{cases} \in \mathcal{A}$ and $\varsigma_i = \begin{cases} \alpha_i \text{ if } i \in \{1, \ldots, n\} \\ \beta_i \text{ if } i \in \{n+1, \ldots, n+m\} \end{cases} \in \mathbb{R}$ which proves that $f + g \in \mathcal{A}$*
3. *Let $f = \sum_{i=1}^{n} \alpha_i \cdot \mathcal{X}_{A_i}$ with $\{\alpha_i\}_{i\in\{1,\ldots,n\}} \Rightarrow \{c \cdot a_i\}_{i\in\{1,\ldots,n\}} \subseteq \mathbb{R}$ and $\{A_i\}_{i\in\{1,\ldots,n\}} \subseteq \mathcal{A}$ with $\mu(A_i) < \infty$ we have that $c \cdot f = \sum_{i=1}^{n} (c \cdot \alpha_i) \cdot \mathcal{X}_{A_i}$ which proves that $c \cdot f \in \mathcal{S}[X, \mathcal{A}]$.* □

The following lemma will allows us to examine the relation between different representations of a simple function.

Lemma 18.116. *Let $\langle X, \mathcal{A}, \mu\rangle$, $n \in \mathbb{N}$ and $\{A_i\}_{i\in\{1,\ldots,n\}} \subseteq \mathcal{A}$ a finite family of sets in $\mathcal{A}$ such that $\forall i \in \mu(A_i) < \infty$. Then there exists a $m \in \mathbb{N}$ and $\{B_i\}_{i\in\{1,\ldots,m\}} \subseteq \mathcal{A}$ such that $\mu(B_i) < \infty$, $\forall i, j \in \{1, \ldots, m\}$ with $i \neq j$ we have $B_i \bigcap B_j = \emptyset$ and $\forall i \in \{1, \ldots, n\}$ there exists a $I_i \subseteq \{1, \ldots, m\}$ such that $A_i = \bigsqcup_{j\in I_i} B_j$.*

Proof. We prove this by induction so let $\mathcal{N} = \{n \in \mathbb{N} | \forall \{A_i\}_{i\in\{1,\ldots,n\}} \subseteq \mathcal{A}$ there exists a pairwise disjoint $\{B_i\}_{i\in\{1,\ldots,m\}} \subseteq \mathcal{A}$ such that $\mu(B_i) < \infty$ and $\forall i \in \{1, \ldots, n\}$ there $\exists I_i \subseteq \{1, \ldots, m\}$ such that $A_i = \bigsqcup_{j\in I_i} B_j\}$ then we have

$1 \in \mathcal{N}$. then for $\{A_i\}_{i\in\{1,\ldots,1\}} \subseteq \mathcal{A}$ take $\{B\}_{i\in\{1,\ldots,1\}}$ by $B_1 = A_1$ then clearly $\{B_i\}_{i\in\{1,\ldots,1\}} \subseteq \mathcal{A}$, $\mu(B_i) = \mu(A_i) < \infty$ and $\forall i \in \{1, \ldots, 1\}$ we have $A_i = A_1 = B_1 = \bigsqcup_{i\in\{1,\ldots,1\}} B_i$. this proves that $1 \in \mathcal{N}$

$n \in \mathcal{N} \Rightarrow n + 1 \in \mathcal{N}$. take $\{A_i\}_{i\in\{1,\ldots,n+1\}}$ then as $n \in \mathcal{N}$ we find a $k \in \mathbb{N}$ and a pairwise disjoint $\{C_i\}_{i\in\{1,\ldots.k\}} \subseteq \mathcal{A}$ such that $\mu(C_i) < \infty$ and

$$\forall i \in \{1, \ldots, n\} \text{ there exists a } I_i \subseteq \{1, \ldots, k\} \text{ such that } A_i = \bigsqcup_{\mathrm{ij}\in I_i} C_j \tag{18.204}$$

Define now $\{E_i\}_{i\in\{1,\ldots,k\}}$ by $E_i = C_i \backslash A_{n+1}$, $\{F_i\}_{i\in\{1,\ldots,k\}}$ by $F_i = C_i \bigcap A_{n+1}$ and $G = A_{n+1} \backslash \bigsqcup_{i\in\{1,\ldots,k\}} C_i$. Using the properties of a σ-algebra (see 18.8) it follows that

$$\{E_i\}_{i\in\{1,\ldots,k\}} \subseteq \mathcal{A}, \{F_i\}_{i\in\{1,\ldots,k\}} \subseteq \mathcal{A} \text{ and } G \in \mathcal{A} \tag{18.205}$$

Further $\forall i \in \{1,\ldots,k\}$ we have $\mu(E_i) \leqslant \mu(C_i) < \infty$ [$E_i \subseteq C_i$ and 18.20], $\mu(F_i) \leqslant \mu(C_i) < \infty$ [$F_i \subseteq C_i$ and 18.20] and $\mu(G) < \mu(A_{n+1}) < \infty$ [$D \subseteq A_{n+1}$ and 18.20]. So we have

$$\forall i \in \{1,\ldots,k\}\ \mu(E_i) < \infty, \mu(F_i) < \infty \text{ and } \mu(G) < \infty \tag{18.206}$$

Define now

$$\{B_i\}_{i\in\{1\ldots,2\cdot k+1\}} \text{ by} \begin{cases} E_i \text{ if } i \in \{1,\ldots,k\} \\ F_{i-k} \text{ if } i \in \{k+1,\ldots,2\cdot k\} \\ G \text{ if } i = 2\cdot k+1 \end{cases} \tag{18.207}$$

Let now $i, j \in \{1,\ldots,2\cdot k+1\}$ with $i \neq j$, then we may assume that $i < j$ [otherwise exchange i and j] and we have the the following cases to consider

$\boldsymbol{i = n+1 \wedge k+1 \leqslant j \leqslant 2\cdot k.}$ then

$$\begin{aligned} B_i \bigcap B_j &= G \bigcap F_{j-k} \\ &= \left(A_{n+1} \backslash \bigsqcup_{i\in\{1,\ldots,k\}} C_i\right) \bigcap C_{j-k} \bigcap A_{n+1} \\ &= \emptyset \end{aligned}$$

$\boldsymbol{i = n+1 \wedge 1 \leqslant j \leqslant k.}$ then

$$\begin{aligned} B_i \bigcap B_j &= G \bigcap E_j \\ &= \left(A_{n+1} \backslash \bigsqcup_{i\in\{1,\ldots,k\}} C_i\right) \bigcap (C_j \backslash A_{n+1}) \\ &= \emptyset \end{aligned}$$

$\boldsymbol{k+1 \leqslant j < i \leqslant 2\cdot k.}$ then

$$\begin{aligned} B_i \bigcap B_j &= F_{i-k} \bigcap F_{j-k} \\ &= (C_{i-k} \bigcap A_{n+1}) \bigcap (C_{j-k} \bigcap A_{n+1}) \\ &\underset{i-k\neq j-k}{=} \emptyset \end{aligned}$$

$\boldsymbol{1 \leqslant j \leqslant k \wedge k+1 \leqslant i \leqslant 2\cdot k.}$ then

$$\begin{aligned} B_i \bigcap B_j &= F_{i-k} \bigcap E_j \\ &= (C_{i-k} \bigcap A_{n+1}) \bigcap (C_j \backslash A_{n+1}) \\ &\underset{i-k\neq j-k}{=} \emptyset \end{aligned}$$

$1 \leqslant j < i \leqslant k$. then

$$\begin{array}{rcl} B_i \bigcap B_j & = & E_i \bigcap E_j \\ & = & (C_i \setminus A_{n+1}) \bigcap (C_j \setminus A_{n+1}) \\ & \underset{i \neq j}{=} & \emptyset \end{array}$$

Using the above we have proved that

$$\forall i, j \in \{1, \ldots, 2 \cdot k + 1\} \text{ with } i \neq j \text{ we have } B_i \bigcap B_j = \emptyset \tag{18.208}$$

Next let $i \in \{1, \ldots, n\}$ we have

$$\begin{array}{rcl} A_i & = & (A_i \setminus A_{n+1}) \bigsqcup (A_i \bigcap A_{n+1}) \\ & \underset{(18.204)}{=} & \left(\left(\bigsqcup_{j \in I_i} C_j\right) \setminus A_{n+1}\right) \bigcup \left(\left(\bigsqcup_{j \in I_i} C_j\right) \bigcap A_{n+1}\right) \\ & \underset{1.108}{=} & \left(\bigsqcup_{j \in I_i} (C_j \setminus A_{n+1})\right) \bigcup \left(\left(\bigsqcup_{j \in I_i} C_j\right) \bigcap A_{n+1}\right) \\ & \underset{1.107}{=} & \left(\bigsqcup_{j \in I_i} (C_j \setminus A_{n+1})\right) \bigcup \left(\bigsqcup_{j \in I_i} (C_j \bigcap A_{n+1})\right) \\ & = & \left(\bigsqcup_{j \in I_i} B_j\right) \bigcup \left(\bigsqcup_{j \in I_i} B_{j+k}\right) \\ & = & \bigsqcup_{j \in I_i \bigcup \{l+k | l \in I_i\}} B_j \\ & = & \bigsqcup_{j \in J_i} B_j \text{ where } J_i = \bigcup_{j \in I_i \bigcup \{l+k | l \in I_i\}} \subseteq \{1, \ldots, 2 \cdot k + 1\} \end{array} \tag{18.209}$$

Further

$$\begin{array}{rcl} A_{n+1} & = & \left(A_{n+1} \bigcap \left(\bigsqcup_{i \in \{1, \ldots, k\}} C_i\right)\right) \bigsqcup \left(A_{n+1} \setminus \left(\bigsqcup_{i \in \{1, \ldots, k\}} C_i\right)\right) \\ & \underset{1.107}{=} & \bigsqcup_{i \in \{1, \ldots, k\}} (A_{n+1} \bigcap C_i) \bigsqcup \left(A_{n+1} \setminus \left(\bigsqcup_{i \in \{1, \ldots, k\}} C_i\right)\right) \\ & = & \left(\bigsqcup_{i \in \{1, \ldots, k\}} B_{i+k}\right) \bigsqcup B_{2 \cdot k+1} \\ & = & \bigsqcup_{i \in \{1, \ldots, k\} \bigcup \{2 \cdot k+1\}} B_i \\ & = & \bigsqcup_{j \in J_{n+1}} B_j \text{ where } J_{n+1} = \{1, \ldots, k\} \bigcup \{2 \cdot k + 1\} \end{array} \tag{18.210}$$

Finally (18.205),(18.207),(18.208),(18.209) and (18.210) proves that $n+1 \in \mathcal{N}$. □

We use the above to prove the following lemma that gives a way to deal with the different representations of a simple function.

Lemma 18.117. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space then we have*

1. *If $f \in \mathcal{S}[X, \mathcal{A}]$ then there exists a **pairwise disjoint** family $\{A_i\}_{i \in \{1,\ldots,n\}} \subseteq \mathcal{A}$ with $\mu(A_i) < \infty, \{\alpha_i\}_{i \in \{1,\ldots,n\}} \subseteq \mathbb{R}$ such that $f = \sum_{i=1}^n \alpha_i \cdot \mathcal{X}_{|A_i}$.*
2. *If $f \in \mathcal{S}[X, \mathcal{A}]$ with $\forall x \in X$ $0 \leqslant f(x)$ we have for every $\{B_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$ with $\mu(B_i) < \infty$ and $\{\beta_i\}_{i \in \mathbb{N}} \subseteq \mathbb{R}$ such that $f = \sum_{i=1}^n \beta_i \cdot \mathcal{X}_{B_i}$ then $0 \leqslant \sum_{i=1}^n \beta_i \cdot \mu(B_i)$.*

Proof. As $f \in \mathcal{S}[X, \mathcal{A}]$ hen there exists a family $\{B_i\}_{i \in \{1,\ldots,n\}} \subseteq \mathcal{A}$ with $\mu(B_i) < \infty$,$\{\beta_i\}_{i \in \{1,\ldots,n\}} \subseteq \mathbb{R}$ such that $f = \sum_{i=1}^n \beta_i \cdot \mathcal{X}_{B_i}$. Using 18.116 there exists a **pairwise disjoint** family $\{A_i\}_{i \in \{1,\ldots,m\}} \subseteq \mathcal{A}$ with $\mu(A_i) < \infty$ such that $\forall i \in \{1, \ldots, n\}$ there exists a $I_i \subseteq \{1, \ldots, m\}$ such that $B_i = \bigsqcup_{j \in I_i} A_j$. Define now $(i, j) \in \bigcup_{i \in \{1,\ldots,n\}} \{(i, j) | j \in I_i\}$ $\zeta_{i,j} = \left\{ \begin{smallmatrix} 1 \text{ if } j \in I_i \\ 0 \text{ otherwise} \end{smallmatrix} \right.$ so that $\beta_i \cdot \zeta_{i,j} = \left\{ \begin{smallmatrix} \beta_i \text{ if } j \in I_i \\ 0 \end{smallmatrix} \right.$ toherwise then we have $\forall x \in X$

$$
\begin{aligned}
f(x) &= \sum_{i=1}^n \beta_i \cdot \mathcal{X}_{B_i}(x) \\
&= \sum_{i=1}^n \beta_i \cdot \mathcal{X}_{\bigsqcup_{j \in I_i} A_j}(x) \\
&\underset{18.110}{=} \sum_{i=1}^n \beta_i \cdot \left(\sum_{j \in I_i} \mathcal{X}_{A_j}(x) \right) \\
&= \sum_{i=1}^n \left(\sum_{j \in I_i} \beta_i \cdot \mathcal{X}_{A_j}(x) \right) \\
&= \sum_{i=1}^n \left(\sum_{j \in I_i} \beta_i \cdot \zeta_{i,j} \cdot \mathcal{X}_{A_j}(x) \right) \\
&= \sum_{i \in \{1,\ldots,n\}} \left(\sum_{j \in I_i} \beta \cdot \zeta_{i,j} \cdot \mathcal{X}_{A_j}(x) \right) \\
&= \sum_{i \in \{1,\ldots,n\}} \left(\left(\sum_{j \in I_i} \beta \cdot \zeta_{i,j} \cdot \mathcal{X}_{A_j}(x) \right) + 0 \right) \\
&= \sum_{i \in \{1,\ldots,n\}} \left(\left(\sum_{j \in I_i} \beta_i . \zeta_{i,j} \cdot \mathcal{X}_{A_j}(x) \right) + \left(\sum_{j \in \{1,\ldots,m\} \setminus I_i} \beta_i \cdot \xi_{i,j} \cdot \mathcal{X}_{A_j}(x) \right) \right) \\
&= \sum_{i \in \{1,\ldots,n\}} \left(\sum_{j \in \{1,\ldots,m\}} \beta_i \cdot \xi_{i,j} \cdot \mathcal{X}_{A_j}(x) \right) \\
&\underset{10.48}{=} \sum_{j \in \{1,\ldots,m\}} \left(\sum_{i \in \{1,\ldots,n\}} \beta_i \cdot \xi_{i,j} \cdot \mathcal{X}_{A_j}(x) \right) \\
&= \sum_{j \in \{1,\ldots,m\}} \left(\sum_{i \in \{1,\ldots,n\}} \beta_i \cdot \xi_{i,j} \right) \cdot \mathcal{X}_{A_j}(x) \\
&= \sum_{j \in \{1,\ldots,m\}} \alpha_j \cdot \mathcal{X}_{A_j}(x) \text{ where } \alpha_j = \sum_{i \in \{1,\ldots,n\}} \beta_i \cdot \xi_{i,j}
\end{aligned}
$$

proving that

$$f = \sum_{j=1}^{m} \alpha_j \cdot \mathcal{X}_{A_j} \text{ where } \alpha_j = \sum_{i \in \{1,\ldots,n\}} \beta_i \cdot \xi_{i,j} = \sum_{i=1}^{n} \beta_i \cdot \varepsilon_{i,j} \tag{18.211}$$

which as $\{A_i\}_{i\in\{1,\ldots,m\}}$ is pairwise disjoint with $\mu(A_i) < \infty$ proves (1) of the lemma. For the second part assume that we have that $\forall x \in X\ f(x) \geqslant 0$. Then for $i \in \{1,\ldots, m\}$ we have for A_i either

- $\boldsymbol{A_i = \emptyset}$**.** then $\alpha_i \cdot \mu(A_i) = 0 \geqslant 0$
- $\boldsymbol{A_i \neq \emptyset}$**.** then $\forall x \in A_i$ such that $0 \leqslant f(x) = \sum_{j=1}^{m} \alpha_j \cdot \mathcal{X}_{A_i}(X) = \alpha_j$ [as $\{A_i\}_{i\in\{1,\ldots,m\}}$ are pairwise disjoint] proving that $\alpha_i \cdot \mu(A_i) \geqslant 0$.

giving

$$\forall i \in \{1,\ldots,m\} \quad 0 \leqslant \alpha_i \cdot \mu(A_i) \tag{18.212}$$

Now

$$\begin{aligned}
\sum_{i=1}^{n} \beta_i \cdot \mu(B_i) &= \sum_{i=1}^{n} \beta_i \cdot \mu\left(\bigsqcup_{j\in I_i} A_j\right) \\
&= \sum_{i=1}^{n} \beta_i \cdot \left(\sum_{j\in I_i} \mu(A_j)\right) \\
&= \sum_{i=1}^{n} \beta_i \cdot \left(\sum_{j\in I_i} \mu(A_j) + 0\right) \\
&= \sum_{i=1}^{n} \beta_i \cdot \left(\sum_{j\in I_i} \zeta_{i,j} \cdot \mu(A_j) + \sum_{j\in\{1,\ldots,m\}\setminus I_j} \zeta_{i,j} \cdot \mu(A_j)\right) \\
&= \sum_{i=1}^{n} \beta_i \cdot \sum_{j\in\{1\ldots,m\}} \zeta_{i,j} \cdot \mu(A_j) \\
&= \sum_{i\in\{1,\ldots,n\}} \left(\sum_{j\in\{1,\ldots,m\}} \beta_i \cdot \zeta_{i,j} \cdot \mu(A_j)\right) \\
&\underset{10.48}{=} \sum_{j\in\{1,\ldots,m\}} \left(\sum_{i\in\{1,\ldots,n\}} \beta_i \cdot \zeta_{i,j} \cdot \mu(A_j)\right) \\
&= \sum_{j\in\{1,\ldots,m\}} \left(\sum_{i\in\{1,\ldots,n\}} \beta_i \cdot \zeta_{i,j}\right) \cdot \mu(A_j) \\
&\underset{(18.211)}{=} \sum_{j\in\{1,\ldots m\}} \alpha_j \cdot \mu(A_j) \\
&\geqslant_{(18.212)} 0
\end{aligned}$$

proving that

$$\text{if } \forall x \in X \text{ we have } 0 \leqslant f(x) \text{ then } 0 \leqslant \sum_{i=1}^{n} \beta_i \cdot \mu(B_i)$$

which proves (2) □

Corollary 18.118. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space then if for $f \in \mathcal{S}[X, \mathcal{A}]$ we have $f = \sum_{i=1}^{n} \alpha_i \cdot \mathcal{X}_{A_i}$ and $f = \sum_{i=1}^{m} \beta_i \cdot \mathcal{X}_{B_i}$ where $\{A_i\}_{i\in\{1,\ldots,n\}}, \{B_i\}_{i\in\{1,\ldots,m\}} \subseteq \mathcal{A}$ with $\forall i \in \{1, \ldots, n\}, \forall j \in \{1, \ldots, m\}$ $\mu(A_i) < \infty \wedge \mu(B_j) < \infty$ and $\{\alpha_i\}_{i\in\{1,\ldots,n\}}$, $\{\beta_i\}_{i\in\{1,\ldots,m\}} \subseteq \mathbb{R}$ we have that $\sum_{i=1}^{n} \alpha_i \cdot \mu(A_i) = \sum_{i=1}^{m} \beta_i \cdot \mu(B_i)$*

Proof. Define $\{\zeta_i\}_{i\in\{1,\ldots,n+m\}}$ by $\zeta_i = \left\{ \begin{smallmatrix} \alpha_i \text{ if } i \in \{1, \ldots, n\} \\ -\beta_{i-n} \text{ if } i \in \{n+1, \ldots, n+m\} \end{smallmatrix} \right.$, $\gamma_i = \left\{ \begin{smallmatrix} -\alpha_i \text{ if } i \in \{1, \ldots, n\} \\ \beta_{i-m} \text{ if } i \in \{n+1, \ldots, n+m\} \end{smallmatrix} \right.$, $\{C_i\}_{i\in\{1,\ldots,n+m\}}$ by $C_i = \left\{ \begin{smallmatrix} A_i \text{ if } i \in \{1, \ldots, n\} \\ B_i \text{ if } i \in \{n+1, \ldots, n+m\} \end{smallmatrix} \right.$ then we have $\forall x \in X$

$$\sum_{i=1}^{n+m} \zeta_i \cdot \mathcal{X}_{C_i}(x) = \sum_{i=1}^{n} \alpha_i \cdot \mathcal{X}_{A_i}(x) - \sum_{i=1}^{m} \beta_i \cdot \mathcal{X}_{B_i}(x) = f(x) - f(x) = 0 \geqslant 0$$

so using 18.117 we have that $0 \leqslant \sum_{i=1}^{n+m} \zeta_i \cdot \mu(C_i) = \sum_{i=1}^{n} \alpha_i \cdot \mu(A_i) - \sum_{i=1}^{m} \beta_i \cdot \mu(B_i)$ giving

$$\sum_{i=1}^{m} \beta_i \cdot \mu(B_i) \leqslant \sum_{i=1}^{n} \alpha_i \cdot \mu(A_i) \tag{18.213}$$

Further

$$\sum_{i=1}^{n+m} \gamma_i \cdot \mathcal{X}_{C_i}(x) = \sum_{i=1}^{m} \beta_{i_i} \cdot \mathcal{X}_{B_i}(x) - \sum_{i=1}^{n} \alpha_i \cdot \mathcal{X}_{A_i}(x) = f(x) - f(x) = 0 \geqslant 0$$

so using 18.117 we have that $0 \leqslant \sum_{i=1}^{n+m} \gamma_i \cdot \mu(C_i) = \sum_{i=1}^{m} \beta_i \cdot \mu(B_i) - \sum_{i=1}^{n} \alpha_i \cdot \mu(A_i)$ which proves that $\sum_{i=1}^{n} \alpha_i \cdot \mu(A_i) \leqslant \sum_{i=1}^{m} \beta_i \cdot \mu(B_i)$. Combining this with (18.213) proves

$$\sum_{i=1}^{n} \alpha_i \cdot \mu(A_i) = \sum_{i=1}^{m} \beta_i \cdot \mu(B_i)$$

□

The above corollary proves that the following definition is well defined

Definition 18.119. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space then for all $f \in \mathcal{S}[X, \mathcal{A}]$ we define the integral $\int^{+} f d\mu$ of the simple function f by*

$$\int^{S} f d\mu = \sum_{i=1}^{n} \alpha_i \cdot \mu(A_i)$$

where $f = \sum_{i=1}^{n} \alpha_i \cdot \mathcal{X}_{A_i}, \{A_i\}_{i\in\{1,\ldots,n\}} \subseteq \mathcal{A}$ with $\mu(A_i) < \infty$ and $\{\alpha_i\}_{i\in\{1,\ldots,n\}} \subseteq \mathbb{R}$

Proposition 18.120. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space then we have*

1. *$\forall f, g \in \mathcal{S}[X, \mathcal{A}]$ that $\int^{S} (f+g) d\mu = \int^{S} f d\mu + \int^{S} g d\mu$ (note that $f + g \in \mathcal{S}[X, \mathcal{A}]$ [see 18.115]*
2. *$\forall f \in \mathcal{S}[X, \mathcal{A}]$ and $c \in \mathbb{R}$ then $\int^{S} (c \cdot f) d\mu = c \cdot \int^{S} f d\mu$*
3. *If $f, g \in \mathcal{S}[X, \mathcal{A}]$ such that $f \leqslant g$ [meaning that $\forall x \in X$ $f(x) \leqslant g(x)$] we have $\int^{S} f d\mu \leqslant \int^{S} g d\mu$*

Proof.

1. Let $f=\sum_{i=1}^{n}\alpha_i\cdot\mathcal{X}_{A_i}$ and $g=\sum_{i=1}^{m}\beta_i\cdot\mathcal{X}_{B_i}$ be representations of f and g then if we define $\{\zeta_i\}_{i\in\{1m,\ldots,\mathrm{mn}+m\}}$ by $\zeta_i=\left\{\begin{smallmatrix}\alpha_i\, i\in\{1,\ldots,n\}\\ \beta_{i-n}\,\text{if}\, i\in\{n+1,\ldots,n+m\}\end{smallmatrix}\right.$ and $\{C_i\}_{i\in\{1,\ldots,n+m\}}$ by $C_i=\left\{\begin{smallmatrix}A_i\,\text{if}\, i\in\{1,\ldots,n\}\\ B_{i-n}\,\text{if}\, i\in\{n+1,\ldots,n+m\}\end{smallmatrix}\right.$ then $\sum_{i=1}^{n+m}\zeta_i\cdot\mathcal{X}_{C_i}$ is a representation of $f+g$ hence $\int^S(f+g)d\mu=\sum_{i=1}^{n+m}\zeta_i\cdot\mu(C_i)=\sum_{i=1}^{n}\alpha_i\cdot\mu(A_i)+\sum_{i=1}^{m}\beta_i\cdot\mu(B_i)=\int^S fd\mu+\int^S gd\mu$
2. Let $f=\sum_{i=1}^{n}\alpha_i\cdot\mathcal{X}_{A_i}$ be a representation of f then $c\cdot f$ has $\sum_{i=1}^{n}(c\cdot\alpha_i)\cdot\mathcal{X}_{A_i}$ as a representation. So that $\int^S(c\cdot f)d\mu=\sum_{i=1}^{n}(c\cdot\alpha_i)\cdot\mu(A_i)=c\cdot\sum_{i=1}^{n}\alpha_i\cdot\mu(A_i)=c\cdot\int^S fd\mu$
3. Let $f=\sum_{i=1}^{n}\alpha_i\cdot\mathcal{X}_{A_i}$ and $g=\sum_{i=1}^{m}\beta_i\cdot\mathcal{X}_{B_i}$ are representations of f and g then if we define $\{\zeta_i\}_{i\in\{1,\ldots,n\}}$ by $\zeta_i=\left\{\begin{smallmatrix}\beta_i\,\text{if}\, i\in\{1,\ldots,n\}\\ -\alpha_{i-n}\,\text{if}\, i\in\{n+1,\ldots,n+m\}\end{smallmatrix}\right.$ and $\{C_i\}_{i\in\{1,\ldots,n\}}$ by $C_i=\left\{\begin{smallmatrix}A_i\,\text{if}\, i\in\{1,\ldots,n\}\\ B_{i-n}\,\text{if}\, i\in\{n+1,\ldots,n+m\}\end{smallmatrix}\right.$ we have that $g-f$ has as representation $\sum_{i=1}^{n+m}\zeta_i\cdot\mathcal{X}_{C_i}$. So $\forall x\in X$ we have $0\leqslant g(x)-f(x)=(g-f)(x)$ so using 18.117 we have that

$$0\leqslant\sum_{i=1}^{n+m}\zeta_i\cdot\mu(C_i)=\sum_{i=1}^{m}\beta_i\cdot\mu(B_i)-\sum_{i=1}^{n}\alpha_i\cdot\mu(A_i)=\int^S gd\mu-\int^S fd\mu$$

proving that

$$\int^S fd\mu\leqslant\int^S gd\mu$$

□

Lemma 18.121. *Let $\langle X,\mathcal{A},\mu\rangle$ be a measure spaces and $f\in\mathcal{S}[X,\mathcal{A}]$ then we have*

1. $\{x\in X|f(x)\neq 0\}\in\mathcal{A}$ *and* $\mu(\{x\in X|f(x)\neq 0\})<\infty$
2. $\exists M\in\mathbb{R}$ *such that* $\forall x\in X$ *we have* $f(x)\leqslant M$ *[in other words f is bounded above]*

Proof.

1. Take $A=\{x\in X|f(x\neq 0)\}$ then we have

$$\begin{aligned}x\in A &\Leftrightarrow x\in X\wedge f(x)\neq 0\\ &\Leftrightarrow x\in X\wedge(f(x)<0\vee 0<f(x))\\ &\Leftrightarrow x\in\{f<0\}\vee x\in\{0<f\}\\ &\Leftrightarrow x\in\{f<0\}\bigcup\{0<f\}\end{aligned}$$

proving that $A=\{f<0\}\bigcup\{0<f\}\in\mathcal{A}$ [as $f:X\to\mathbb{R}$]. Let $\sum_{i=1}^{n}\alpha_i\cdot\mathcal{X}_{A_i}$be a representation of the simple function f then if $f(x)\neq 0$ there must exists a $i\in\{1,\ldots,n\}$ such that $x\in A_i$ [otherwise $\forall i\in\{1,\ldots,n\}$ we have $x\notin A_i=\mathcal{X}_{A_i}(x)=0$ proving that $f(x)=0$ a contradiction]. Hence we have that $A\subseteq\bigcup_{i\in\{1,\ldots,n\}}A_i$ so using 18.20 we have that $\mu(A)\leqslant\mu\left(\bigcup_{i\in\{1,\ldots,n\}}A_i\right)<\infty$ [as $\forall i\in\{1,\ldots,n\}$ $\mu(A_i)<\infty$]. Which proves the (1).

2. Let $\sum_{i=1}^{n} \alpha_\iota \cdot \mathcal{X}_{|A_i}$ be a representation of the simple function f then we have for $x \in X$ that $f(x) = \sum_{i=1}^{n} \alpha_i \cdot \mathcal{X}_{A_i}(x) \leqslant \sum_{i=1}^{n} |\alpha_i \cdot \mathcal{X}_{A_i}(x)| = \sum_{i=1}^{n} |\alpha_i| \cdot |\mathcal{X}_{A_i}(x)| \leqslant \sum_{i=1}^{n} |\alpha_i|$, so taking $M = \sum_{i=1}^{n} |\alpha_i|$ proves (2).

□

Lemma 18.122. *Let $\langle X, \mathcal{A}, \mu \rangle$ be a measure space, $\{f_i\}_{i\in\mathbb{N}} \subseteq \mathcal{S}[X, \mathcal{A}]$ such that $\forall i \in \mathbb{N}$ we have $f_i \leqslant f_{i+1}$ and $f \in \mathcal{S}[X, \mathcal{A}]$ such that $f(x) \leqslant \sup(\{f_i(x) | i \in \mathbb{N}\})$ for almost every $x \in X$ [see Algorithm 18.28] then $\int^S f d\mu \leqslant \sup\left(\left\{\int^S f_i d\mu | i \in \mathbb{N}\right\}\right)$*

Proof. Note that as f, f_0 are simple functions we have that $f - f_0$ is a simple function (see 18.115). If we define then $H = \{x \in X | (f - f_0)(x) \neq 0\}$ it follows from 18.121 that

$$\mu(H) < \infty \tag{18.214}$$

and

$$\exists M \in \mathbb{R} \text{ such that } \forall x \in X \text{ we have } (f - f_0)(x) \leqslant M \text{ where } 0 \leqslant M \tag{18.215}$$

Take now $\delta > 0$ and take

$$\varepsilon = \frac{\delta}{1 + \mu(H) + M} \tag{18.216}$$

Take $n \in \mathbb{N}$ then $f - f_n$ is a simple function hence $\mathcal{A}$-measurable [see 18.115] so that $\{\varepsilon \leqslant f - f_n\} \in \mathcal{A}$ [see 18.98]. Further if $x \in \{\varepsilon \leqslant f - f_{n+1}\}$ then $\varepsilon \leqslant (f - f_{n+1})(x) = f(x) - f_{n+1}(x) \leqslant f(x) - f_n(x)$ proving that $x \in \{\varepsilon \leqslant f - f_n\}$. to summarize we have

$$\forall n \in \mathbb{N}\ \{\varepsilon \leqslant f - f_n\} \in \mathcal{A} \text{ and } \{\varepsilon \leqslant f - f_{n+1}\} \subseteq \{\varepsilon \leqslant f - f_n\} \tag{18.217}$$

Let $x \in \bigcap_{n\in\mathbb{N}} \{\varepsilon \leqslant f - f_n\}$ then $\forall n \in \mathbb{N}$ we have $\varepsilon \leqslant (f - f_n)(x) = f(x) - f_n(x)$ hence $f_n(x) \leqslant f(x) - \varepsilon$ so that $\sup(\{f_n(x) | n \in \mathbb{N}\}) \leqslant f(x) - \varepsilon$ or $\sup(\{f_n(x) | n \in \mathbb{N}\}) + \varepsilon \leqslant f(x)$ so that $\sup(\{f_n(x) | n \in \mathbb{N}\}) < f(x)$ giving

$$\bigcap_{n\in\mathbb{N}} \{\varepsilon \leqslant f - f_n\} \subseteq \{x \in X | \sup(\{f_n(x) | n \in \mathbb{N}\}) < f(x)\} \tag{18.218}$$

As by the hypothesis $f(x) \leqslant \sup(\{f_n(x)\})$ a.e. $\{x \in X | f(x) \nleqslant \sup(\{f_n | n \in \mathbb{N}\})\}$ is negligible hence there exists a $B \in \mathcal{A}$ with $\mu(B) = 0$ such that $\{x \in X | \sup(\{f_n | n \in \mathbb{N}\}) < f(x)\} = \{x \in X | f(x) \nleqslant \sup(\{f_n | n \in \mathbb{N}\})\} \subseteq B$. Using 18.8 (1) and (18.217) we have that $\bigcap_{n\in\mathbb{N}} \{\varepsilon \leqslant f - f_n\} \in \mathcal{A}$, further $0 \leqslant \mu(\bigcap_{n\in\mathbb{N}} \{\varepsilon \leqslant f - f_n\}) \leqslant \mu(B) = 0$, so we have

$$\mu\left(\bigcap_{n\in\mathbb{N}} \{\varepsilon \leqslant f - f_n\}\right) = 0 \tag{18.219}$$

Further as $x \in \{\varepsilon \leqslant f - f_0\} \Rightarrow 0 < \varepsilon \leqslant (f - f_0)(x) \Rightarrow x \in H$ proving $\{\varepsilon \leqslant f - f_0\}$ we have using (18.214) that

$$\mu(\{\varepsilon \leqslant f - f_0\}) < \infty \tag{18.220}$$

(18.217),(18.219) and (18.220) allows us to apply 18.20 (6) giving

$$\inf(\{\mu(\{\varepsilon \leqslant f - f_n\}) | n \in \mathbb{N}\}) = 0 \tag{18.221}$$

As $0<\varepsilon$ we have by the definition of a infinum and the above that

$$\exists n\in\mathbb{N}\,\text{such that}\,\mu(\{\varepsilon\leqslant f-f_n\})<\varepsilon \tag{18.222}$$

Define the simple function $h=\varepsilon\cdot\mathcal{X}_H+M\cdot\mathcal{X}_{\{\varepsilon\leqslant f-f_n\}}$ then as $0<\varepsilon, 0\leqslant M$ we have that $0\leqslant h$. Then for the simple function $g=f_n+h$ we have that $f\leqslant g$ and

$$\begin{aligned}
\int^S f d\mu &\leqslant_{18.120} \int^S g d\mu \\
&\underset{18.120}{=} \int^S f_n d\mu+\int^S h d\mu \\
&= \int^S f_n\, d\mu+\varepsilon\cdot\mu(H)+M\cdot\mu(\mathcal{X}_{\{\varepsilon\leqslant f-f_n\}}) \\
&<_{(18.222)} \int^S f_n d\mu+\varepsilon\cdot\mu(H)+M\cdot\varepsilon \\
&= \int^S f_n\, d\mu+\varepsilon\cdot(\mu(H)+M) \\
&\underset{(18.216)}{=} \int^S f_n\, d\mu+\delta
\end{aligned}$$

As δ is chosen arbitrary we have using 9.56 that $\int^S f d\mu\leqslant\int^S f_n\leqslant\sup\left(\left\{\int^S f_n d\mu|n\in\mathbb{N}\right\}\right)$ proving that

$$\int^S f d\mu\leqslant\sup\left(\left\{\int^S f_n\, d\mu|n\in\mathbb{N}\right\}\right)$$

□

Lemma 18.123. *Let $\langle X,\mathcal{A},\mu\rangle$ be a measure space $\{f_n\}_{n\in\mathbb{N}}\subseteq\mathcal{S}[X,\mathcal{A}]$ such that $\forall n\in\mathbb{N}$ we have $f_n\leqslant f_{n+1}$ then $\lim_{n\rightarrow\infty}\int^S f_n d\mu<\infty$ if and only if $\sup\left(\left\{\int^S f_n d\mu|n\in\mathbb{N}\right\}\right)<\infty$. If the limit or supremum exists then $\lim_{n\rightarrow\infty}\int^S f_n d\mu=\sup\left(\left\{\int^S f_n d\mu|n\in\mathbb{N}\right\}\right)$.*

Proof. Using 18.120 it follows from $\forall n\in\mathbb{N}\;\; f_n\leqslant f_{n+1}$ that $\forall n\in\mathbb{N}\;\int^S f_n d\mu\leqslant\int^S f_n d\mu$ proving that $\left\{\int^S f_n d\mu\right\}_{n\in\mathbb{N}}$ is a increasing sequence of real numbers. the lemma follows then from applying 12.354. □

18.2.2.2 Integral of positive integrable functions

Lemma 18.124. *Let $\langle X,\mathcal{A}\rangle$ be a measurable space and $\{f_n: X\rightarrow\mathbb{R}\}_{n\in\mathbb{N}}\subseteq\mathcal{S}[X,\mathcal{A}]$ a sequence of simple functions [so $\forall n\in\mathbb{N}$ $\mathrm{dom}(f_n)=X$) then*

1. $\mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right)=\{x\in X|\{x_n\}_{n\in\mathbb{N}}\,\text{has}\,a\,\text{limit in}\,\mathbb{R}\}$

2. $\forall x\in\mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right)$ $\left(\lim_{n\rightarrow\infty} f_n\right)(x)$ *is the limit of* $\{f_i(x)\}_{i\in\mathbb{N}}$

3. *If* $\forall n\in\mathbb{N}$ *we have* $f_n\leqslant f_{n+1}$ *then* $\left(\lim_{n\rightarrow\infty} f_i\right)(x)=\sup\left(\{f_i(x)|i\in\mathbb{N}\}\right)$

Proof. This follows from 18.89 and the fact that for simple functions we have $\forall n\in\mathbb{N}\,\mathrm{dom}(f_n)=X$. □

Lemma 18.125. *Let $\langle X,\mathcal{A},\mu\rangle$ be a measure space, $f:X\rightarrow\mathbb{R}$ a partial function such that there exists a sequence $\{f_n\}_{n\in\mathbb{N}}\subseteq\mathcal{S}[X,\mathcal{A}]$ such that*

1. *$\forall n\in\mathbb{N}$ we have $f_n\leqslant f_{n+1}$*
2. *$\forall n\in\mathbb{N}$ $\sup\left(\left\{\int^S f_n d\mu|n\in\mathbb{N}\right\}\right)<\infty$ [or using 18.123 $\lim_{n\rightarrow\infty}\int^S f_n d\mu<\infty$]*
3. *$\lim_{n\rightarrow\infty} f_n(x)=_{a.e.} f(x)$ [see Algorithm 18.29]*

then we have that

$$\sup\left(\left\{\int^S f_n d\mu|n\in\mathbb{N}\right\}\right)=\sup\left(\left\{\int^S g d\mu|g\in\mathcal{S}[X,\mathcal{A}]\text{ such that }g\leqslant_{a.e.} f\right\}\right)<\infty$$

here $g\leqslant_{a.e.} f$ means that $\{x\in\mathrm{dom}(f)|g(x)\leqslant f(x)\}$ is conegligible

Proof. From (3) it follows that

$$E=\left\{x\in\mathrm{dom}(f)\bigcap\mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right)|\lim_{n\rightarrow\infty} f_n(x)=f(x)\right\}\text{ is conegligible}\qquad(18.223)$$

Given $x\in E$ we have as $\{f_n(x)\}_{i\in\mathbb{N}}$ is a increasing sequence using 12.354 that

$$\forall x\in E\quad f(x)=\lim_{n\rightarrow\infty} f_n(x)=\sup\left(\{f_n(x)|n\in\mathbb{N}\}\right)\Rightarrow\forall n\in\mathbb{N}\text{ we have }f_n(x)\leqslant f(x)\qquad(18.224)$$

As E is conegligible by (18.223) it follows that

$$\forall n\in\mathbb{N}\text{ we have }f_n\leqslant_{a.e.} f\qquad(18.225)$$

which proves that $\left\{\int^S f_n d\mu|n\in\mathbb{N}\right\}\subseteq\left\{\int^S g d\mu|g\in\mathcal{S}[X,\mathcal{A}]\text{ such that }g\leqslant_{a.e.} f\right\}$. Hence using 2.171 and the fact that the supremum always exists in $\bar{\mathbb{R}}$ we have

$$\sup\left(\left\{\int^S f_n d\mu|n\in\mathbb{N}\right\}\right)\leqslant\sup\left(\left\{\int^S g d\mu|g\in\mathcal{S}[X,\mathcal{A}]\text{ such that }g\leqslant_{a.e.} f\right\}\right)\leqslant\infty\qquad(18.226)$$

For the opposite inequality take $g\in\mathcal{S}[X,\mathcal{A}]$ such that $g\leqslant_{a.e.} f$ then $D=\{x\in\mathrm{dom}(f)|g(x)\leqslant f(x)\}$ is conegligible, hence $\forall x\in D\bigcap E$ [a conegligible set by 18.27] $g(x)\leqslant f(x)\underset{(18.224)}{=}\sup\left(\{f_n(x)|n\in\mathbb{N}\}\right)$ proving that $g\leqslant_{a.e.}\sup\left(\{f_n(x)|n\in\mathbb{N}\}\right)$, so using 18.122 it follows that $\int^S g d\mu\leqslant\sup\left(\left\{\int^S f_n d\mu|n\in\mathbb{N}\right\}\right)$. Hence $\sup\left(\left\{\int^S g d\mu|g\in\mathcal{S}[X,\mathcal{A}]\text{ such that }g\leqslant_{a.e.} f\right\}\right)\leqslant\sup\left(\left\{\int^S f_n d\mu|n\in\mathbb{N}\right\}\right)<\infty$ [$<\infty$ as stated in the conditions for the lemma] which by (18.226) proves finally

$$\sup\left(\left\{\int^S f_n d\mu|n\in\mathbb{N}\right\}\right)=\sup\left(\left\{\int^S g d\mu|g\in\mathcal{S}[X,\mathcal{A}]\text{ such that }g\leqslant_{a.e.} f\right\}\right)<\infty\ \Box$$

We are now ready to define the concept of integrability on non negative real partial functions.

Definition 18.126. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space then a function $f: X \rightarrow \mathbb{R}$ is a positive integrable partial function iff*

1. $\forall x \in \operatorname{dom}(f)$ *we have* $f(x) \in [0, \infty[$
2. *There exists a sequence* $\{f_n\}_{n \in \mathbb{N}}$ *of simple functions such that*
   a. $\forall n \in \mathbb{N} \quad 0 \leqslant f_n \leqslant f_{n+1}$
   b. $\sup\left(\left\{\int^S f_n d\mu | n \in \mathbb{N}\right\}\right) < \infty$ *[or using 18.123* $\lim_{n \rightarrow \infty} \int^S f_n d\mu < \infty$*]*
   c. $\lim_{n \rightarrow \infty} f_n(x) =_{a.e.} f(x)$

the set of all positive integrable functions is noted as $\mathcal{L}_+[X, \mathcal{A}, \mu]$

Remark 18.127. As by $(2.c)$ $\left\{x \in \operatorname{dom}(f) \bigcap \operatorname{dom}\left(\lim_{n \rightarrow \infty} f_n\right) | \lim_{n \rightarrow \infty} f_n(x) = f(x)\right\}$ is assumed to be conegligible and $\left\{x \in \operatorname{dom}(f) \bigcap \operatorname{dom}\left(\lim_{n \rightarrow \infty} f_n\right) | \lim_{n \rightarrow \infty} f_n(x) = f(x)\right\} \subseteq \operatorname{dom}(f)$ we have by 18.27 that $\operatorname{dom}(f)$ is conegligible.

The following lemma shows how every positive valued partial function can be written as a limit of linear combinations of characteristics functions. This will be used to simplify the conditions for a partial function to be positive integrable because it will allows us too construct the simple functions mentioned in the definition.

Lemma 18.128. *Let X be a set, $f: X \rightarrow \mathbb{R}$ a partial function with* $\operatorname{dom}(f) \subseteq X$ $f(X) \in [0, \infty[$ *then there exists a sequence* $\{f_n\}_{n \in \mathbb{N}}$ *of* ***functions*** *where* $f_n: X \rightarrow \mathbb{R}$ *is defined by* $f_n = \sum_{k=1}^{4^n} \frac{1}{2^n} \cdot \mathcal{X}_{\left\{\frac{k}{2^n} \leqslant f\right\}}$ *such that* $\forall n \in \mathbb{N}$ $0 \leqslant f_n \leqslant f_{n+1}$ *and* $f = \lim_{n \rightarrow \infty} f_n = \sup(\{f_n | n \in \mathbb{N}\})$ *[meaning that* $\forall x \in \operatorname{dom}(f)$ *we have* $f(x) = \lim_{n \rightarrow \infty} f_n(x) = \sup(\{f_n(x) | n \in \infty\})$*].*

Proof. Let $n \in \mathbb{N}$. If $x \in \bigcup_{k \in \{0, \ldots, 4^n - 1\}} \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right[$ we have that $\exists k \in \{0, \ldots, 4^n - 1\}$ such that $0 = \frac{0}{2^n} \leqslant \frac{k}{2^n} \leqslant x < \frac{k+1}{2^n} \leqslant \frac{(4^n - 1) + 1}{2^n} = \frac{4^n}{2^n} = 2^n$ which proves that $x \in [0, 2^n[$ hence

$$\bigcup_{k \in \{0, \ldots, 4^n - 1\}} \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right[\subseteq [0, 2^n[\tag{18.227}$$

If $x \in [0, 2^n[\Rightarrow 0 \leqslant x < 2^n$ define then $A_x = \left\{k \in \{0, \ldots, 4^n - 1\} | \frac{k}{2^n} \leqslant x\right\} \subseteq \{0, \ldots, 4^n - 1\}$. As $\frac{0}{2^n} = 0 \leqslant x$ we have that $0 \in A_x$ proving that $A_x \neq \emptyset$. So $l = \max(A_x)$ exists. The following cases must then be considered for l

$l = 4^n - 1$. then $\frac{l}{2^n} = \frac{4^n - 1}{2^n} \leqslant x$ and as $x < 2^n = \frac{(4^n - 1) + 1}{2^n} = \frac{l+1}{2^n}$ we have that $x \in \left[\frac{l}{2^n}, \frac{l+1}{2^n}\right[\subseteq \bigcup_{k \in \{0, \ldots, 4^n - 1\}} \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right[$

$l<4^n-1$. then $\frac{l}{2^n}\leqslant x$ and by the definition of a maximum for $l+1\in\{0,\ldots,4^n-1\}$ we must have $x<\frac{l+1}{2^n}$ proving that $x\in\left[\frac{l}{2^n},\frac{l+1}{2^n}\right[\subseteq\bigcup_{k\in\{0,\ldots,4^n-1\}}\left[\frac{k}{2^n},\frac{k+1}{2^n}\right[$

so it follows that $[0,2^n[\subseteq\bigcup_{k\in\{0,\ldots,4^n-1\}}\left[\frac{k}{2^n},\frac{k+1}{2^n}\right[$ which together with (18.227) gives

$$[0,2^n[=\bigcup_{k\in\{0,\ldots,4^n-1\}}\left[\frac{k}{2^n},\frac{k+1}{2^n}\right[\tag{18.228}$$

Assume now that $\exists k,l\in\{0,\ldots,4^n-1\}$ with $k\neq l$ then we may assume that $k<l$ [otherwise interchange k with l]. Then $k+1\leqslant l$ if $x\in\left[\frac{k}{2^n},\frac{k+1}{2^n}\right[\bigcap\left[\frac{l}{2^n},\frac{l+1}{2^n}\right[$ then $x<\frac{k+1}{2^n}\leqslant\frac{l}{2^n}\leqslant x$ giving the contradiction $x<x$, hence we must have that $\left[\frac{k}{2^n},\frac{k+1}{2^n}\right[\bigcap\left[\frac{l}{2^n},\frac{l+1}{2^n}\right[=\emptyset$. Combining this with (18.228) gives

$$[0,2^n[=\bigsqcup_{k\in\{0,\ldots,4^n-1\}}\left[\frac{k}{2^n},\frac{k+1}{2^n}\right[\tag{18.229}$$

So for every $x\in[0,2^n[$ there exists a unique $k_n(x)\in\{0,\ldots,4^n-1\}$ such that $x_n(x)\in\left[\frac{k}{2^n},\frac{k+1}{2^n}\right[$, this allows us to define the following function

$$k_n\colon [0,2^n[\rightarrow\{0,\ldots,4^n-1\} \text{ where } k_n(x) \text{ satisifes } x\in\left[\frac{k_n(x)}{2^n},\frac{k_n(x)+1}{2^n}\right[\text{ and } \forall k\in\{0,\ldots,4^n-1\} \text{ we have } x\notin\left[\frac{k}{2^n},\frac{k+1}{2^n}\right[\tag{18.230}$$

Define now

$$f_n\colon X\rightarrow\mathbb{R} \text{ by } \sum_{k=1}^{4^n}\frac{1}{2^n}\cdot\mathcal{X}_{\left\{\frac{k}{2^n}\leqslant f\right\}} \tag{18.231}$$

Take $x\in X$ then for x we have the following possible cases:

$x\notin\mathrm{dom}(f)$. Then $\forall k\in\{0,\ldots,4^n\}$ we have $x\notin\left\{\frac{k}{2^n}\leqslant f(x)\right\}$ so that $f_n(x)=0$

$x\in\mathrm{dom}(f)\wedge 2^n\leqslant f(x)$. then $\forall k\in\{0,\ldots,4^n\}$ we have $\frac{k}{2^n}\leqslant\frac{4^n}{2^n}=2^n\leqslant f(x)=x\in\left\{\frac{k}{2^n}\leqslant f\right\}$ hence

$$\begin{aligned} f_n(x) &= \sum_{k=1}^{4^n}\frac{1}{2^n}\cdot 1 \\ &= \frac{4^n}{2^n} \\ &= 2^n \end{aligned}$$

$x \in \mathrm{dom}(f) \wedge 0 \leqslant f(x) < 2^n$. then by (18.230) we have that $\frac{k_n(f(x))}{2^n} \leqslant f(x) < \frac{k_n(f(x)) + 1}{2^n}$ so that $\forall k \in \{1, \ldots, k_n(f(x))\}$ we have $\frac{k}{2^n} \leqslant \frac{k_n(f(x))}{2^n} \leqslant f(x) \Rightarrow x \in \left\{\frac{k}{2^n} \leqslant f\right\}$ and $\forall k \in \{k_n(f(x)) + 1, \ldots, 4^n\}$ we have $f(x) < \frac{k_n(f(x))}{2^n} \leqslant \frac{k}{2^n} \Rightarrow x \notin \left\{\frac{k}{2^n} \leqslant f\right\}$. Hence

$$
\begin{aligned}
f_n(x) &= \sum_{k=1}^{4^n} \frac{1}{2^n} \cdot \mathcal{X}_{\left\{\frac{k}{2^n} \leqslant f\right\}} \\
&= \sum_{k=1}^{k_n(f(x))} \frac{1}{2^n} \cdot \mathcal{X}_{\left\{\frac{k}{2^n} \leqslant f\right\}} + \sum_{k=k_n(f(x))+1}^{4^n} \frac{1}{2^n} \cdot \mathcal{X}_{\left\{\frac{k}{2^n} \leqslant f\right\}} \\
&= \sum_{k=1}^{k_n(f(x))} \frac{1}{2^n} \cdot 1 + \sum_{k=k_n(f(x))+1}^{4^n} \frac{1}{2^n} \cdot 0 \\
&= \frac{k_n(f(x))}{2^n}
\end{aligned}
$$

To summarize these cases we have for f_n that

$$
\forall x \in X \text{ we have } f_n(x) = \begin{cases} 0 \text{ if } x \in X \setminus \mathrm{dom}(f) \\ 2^n \text{ if } 2^n \leqslant f(x) \\ \frac{k_n(f(x))}{2^n} \text{ if } 0 \leqslant f(x) < 2^n \end{cases} \tag{18.232}
$$

Fix $n \in \mathbb{N}$ and take $x \in X$ then we have

$x \notin \mathrm{dom}(f)$. then $f_n(x) = 0 = f_{n+1}(x) \Rightarrow f_n(x) \leqslant f_{n+1}(x)$

$x \in \mathrm{dom}(f)$. then we must consider the following cases

$0 \leqslant f(x) < 2^n$. then $f_n(x) = \frac{k_n(f(x))}{2^n}$ where $k_n(f(x)) \in \{0, \ldots, 4^n - 1\}$ and $\frac{k_n(f(x))}{2^n} \leqslant f(x) < \frac{k_n(f(x)) + 1}{2^n}$, we can rewrite this as follows $\frac{2 \cdot k_n(f(x))}{2^{n+1}} \leqslant f(x) < \frac{2 \cdot (k_n(f(x)) + 1)}{2^{n+1}}$ giving

$$
\frac{2 \cdot k_n(f(x))}{2^{n+1}} \leqslant f(x) < \frac{(2 \cdot k_n(f(x)) + 1) + 1}{2^{n+1}} \tag{18.233}
$$

then we have either

$f(x) < \frac{2 \cdot k_n(f(x)) + 1}{2^{n+1}}$. then $\frac{2 \cdot k_n(f(x))}{2^{n+1}} \leqslant f(x) < \frac{2 \cdot k_n(f(x)) + 1}{2^{n+1}}$ where $2 \cdot k_n(f(x)) \in \{0, \ldots, 2 \cdot (4^n - 1)\} \subseteq \{0, \ldots, 4^{n+1} - 1\}$ [as $2 \cdot (4^n - 1) = 2 \cdot 4^n - 2 \leqslant 4 \cdot 4^n - 1 = 4^{n+1} - 1$] so that $f_{n+1}(x) = k_{n+1}(f(x)) = \frac{2 \cdot k_n(f(x))}{2^{n+1}} = \frac{k_n(f(x))}{2^n} = k_n(f(x)) = f_n(x)$ proving that $f_n(x) \leqslant f_{n+1}(x)$

$\frac{2 \cdot k_n(f(x)) + 1}{2^{n+1}} \leqslant f(x)$. then $\frac{2 \cdot k_n(f(x)) + 1}{2^{n+1}} \leqslant f(x) < \frac{(2 \cdot k_n(f(x)) + 1) + 1}{2^{n+1}}$ where $2 \cdot k_n(f(x)) + 1 \in \{0, \ldots, 2 \cdot (4^n - 1) + 1\} \subseteq \{0, \ldots, 4^{n+1} - 1\}$ [as $2 \cdot (4^n - 1) + 1 = 2 \cdot 4^n - 1 \leqslant 4 \cdot 4^n - 1 = 4^{n+1} - 1$] so that $f_{n+1}(x) = k_{n+1}(f(x)) = \frac{2 \cdot k_n(f(x)) + 1}{2^{n+1}} > \frac{2 \cdot k_n(f(x))}{2^{n+1}} = \frac{k_n(f(x))}{2^n} = f_n(x)$ proving that $f_n(x) \leqslant f_{n+1}(x)$

So in all cases we have

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

$2^n \leqslant f(x) < 2^{n+1}$. then $f_{n+1}(x) = \frac{k_{n+1}(f(x))}{2^{n+1}}$ where $\frac{k_{n+1}(f(x))}{2^{n+1}} \leqslant f(x) < \frac{k_{n+1}(f(x))+1}{2^{n+1}}$ As $2^n \leqslant f(x)$ we have that

$$\begin{aligned} 2^n < \frac{k_{n+1}(f(x))+1}{2^{n+1}} &\Rightarrow 2^n \cdot 2^{n+1} < k_{n+1}(f(x))+1 \\ &\Rightarrow 2^n \cdot 2^{n+1} \leqslant \\ &\Rightarrow 2^n \leqslant \frac{k_{n+1}(f(x))}{2^{n+1}} \\ &\Rightarrow 2^n \leqslant f_{n+1}(x) \end{aligned}$$

which as $f_{n+1}(x) = 2^n$ proves that

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

$2^{n+1} \leqslant f(x)$. then as $2^n < 2^{n+1}$ we have also $2^n < f(x)$ and thus $f_n(x) = 2^n < 2^{n+1} = f_{n+1}(x)$ which proves that $f_n(x) \leqslant f_{n+1}(x)$

So in all cases we have $f_n(x) \leqslant f_{n+1}(x)$ giving as from (18.232) $0 \leqslant f_n$ that

$$\forall n \in \mathbb{N} \text{ we have } 0 \leqslant f_n \leqslant f_{n+1} \tag{18.234}$$

Let $x \in E$ and take $n \in \mathbb{N}$ then we have either

$0 \leqslant f(x) < 2^n$. then $f_n(x) = \frac{k_n(f(x))}{2^n} \leqslant f(x) < \frac{k_n(f(x))+1}{2^n}$ proving that $f_n(x) \leqslant f(x)$

$2^n \leqslant f(x)$. then $f_n(x) = 2^n \leqslant f(x)$ proving also that $f_n(x) \leqslant f(x)$

So $\forall n \in \mathbb{N} \quad f_n(x) \leqslant f(x)$ proving that

$$\sup(\{f_n(x) | n \in \mathbb{N}\}) \leqslant f(x) \tag{18.235}$$

Take $S = \sup(\{f_n(x)|n \in \mathbb{N}\})$ [which exists in $\mathbb{R}$ because $\mathbb{R}$ is conditionally complete 9.43 and $\{f_n(x)|n \in \mathbb{N}\}$ is bounded above]. Assume that $S = \sup(\{f_n(x)|n \in \mathbb{N}\}) < f(x)$ then $0 < f(x) - S$ and using (9.66) there exists a $N_1 \in \mathbb{N}$ such that $0 < \frac{1}{2^{N_1}} < f(x) - S$. Further using 9.64 there exists $N_2, N_3 \in \mathbb{N}$ such that $f(x) < 2^{N_2}$ and $S < 2^{N_3}$. So if we take $N = \max(N_1, N_2, N_3)$ then we have

$$0 < \frac{1}{2^N} < f(x) - S \wedge f(x) \leqslant 2^N \wedge S < 2^N \tag{18.236}$$

Using the above together with (18.230) gives that $\frac{k_N(f(x))}{2^N} \leqslant f(x) < \frac{k_N(f(x))+1}{2^N}$ and $\frac{k_N(S)}{2^N} \leqslant S < \frac{k_N(S)+1}{2^N}$. We have now the following cases to look at:

$k_n(f(x)) = k_n(S)$. then $f(x) - S < \frac{k_N(f(x))+1}{2^N} - \frac{k_N(S)}{2^N} = \frac{k_N(S)+1}{2^N} - \frac{k_N(S)}{2^N} = \frac{1}{2^N}$ contradicting $\frac{1}{2^N} < f(x) - S$.

$k_n(f(x)) < k_n(S)$. then $k_n(f(x)) + 1 \leqslant k_n(S)$ so that $\frac{k_N(f(x))}{2^N} \leqslant f(x) < \frac{k_N(f(x))+1}{2^N} \leqslant \frac{k_N(S)}{2^N} \leqslant S < \frac{k_N(S)+1}{2^N}$ giving $f(x) < S$ contradicting the assumption that $S < f(x)$

$k_n(S) < k_n(f(x))$. then $k_n(S)+1 \leqslant k_n(f(x))$ so that $S < \frac{k_N(S)+1}{2^N} \leqslant \frac{k_N(f(x))}{2^N} \leqslant f(x) < \frac{k_N(f(x))+1}{2^N}$ or as $f_N(x) \underset{(18.232)}{=} \frac{k_N(f(x))}{2^N}$ that $S = \sup(\{f_n(x)|n \in \mathbb{N}\}) < f_N(x)$ which is a contradiction.

As we have in all cases a contradiction our assumption $S = \sup(\{f_n(x)|n \in \mathbb{N}\}) < f(x)$ is false hence we have using (18.235) that

$$f(x) = \sup(\{f_n(x)|n \in \mathbb{N}\}) \underset{(18.234) \text{ and } 12.354}{=} \lim_{n-<\infty} f_n(x)$$ □

Lemma 18.129. *Let $\langle X, \mathcal{A}, \mu \rangle$ be a measure space and $f: X \rightarrow \mathbb{R}$ a partial function and $E \in \mathcal{A}$ such that $E \subseteq \operatorname{dom}(f)$ then if $f_{|E}$ is $\mathcal{A}$-measurable we have that $\forall a \in \mathbb{R}$*

1. $\{a \leqslant f_{|E}\} \in \mathcal{A}$
2. $\{a < f_{|E}\} \in \mathcal{A}$
3. $\{f_{|E} \leqslant a\} \in \mathcal{A}$
4. $\{f_{|E} < a\} \in \mathcal{A}$

Proof. First note that $\operatorname{dom}(f_{|E}) = E \bigcap \operatorname{dom}(f) \underset{E \subseteq \operatorname{dom}(f)}{=} E$ then

1. As $f_{|E}$ is $\mathcal{A}$-measurable we have by 18.98 that $\{a \leqslant f_{|E}\} \in \mathcal{A}_{|\operatorname{dom}(f_{|E})} = \mathcal{A}_{|E}$ so there exists a $A \in \mathcal{A}$ such that $\{a \leqslant f_{|E}\} = A \bigcap E$. Now as $E \in \mathcal{A}$ we have $A \bigcap E \in \mathcal{A}$ so that $\{a \leqslant f_{|E}\} \in \mathcal{A}$
2. As $f_{|E}$ is $\mathcal{A}$-measurable we have by 18.98 that $\{a < f_{|E}\} \in \mathcal{A}_{|\operatorname{dom}(f_{|E})} = \mathcal{A}_{|E}$ so there exists a $A \in \mathcal{A}$ such that $\{a < f_{|E}\} = A \bigcap E$. Now as $E \in \mathcal{A}$ we have $A \bigcap E \in \mathcal{A}$ so that $\{a < f_{|E}\} \in \mathcal{A}$
3. As $f_{|E}$ is $\mathcal{A}$-measurable we have by 18.98 that $\{f_{|E} \leqslant a\} \in \mathcal{A}_{|\operatorname{dom}(f_{|E})} = \mathcal{A}_{|E}$ so there exists a $A \in \mathcal{A}$ such that $\{f_{|E} \leqslant a\} = A \bigcap E$. Now as $E \in \mathcal{A}$ we have $A \bigcap E \in \mathcal{A}$ so that $\{f_{|E} \leqslant a\} \in \mathcal{A}$
4. As $f_{|E}$ is $\mathcal{A}$-measurable we have by 18.98 that $\{f_{|E} < a\} \in \mathcal{A}_{|\operatorname{dom}(f_{|E})} = \mathcal{A}_{|E}$ so there exists a $A \in \mathcal{A}$ such that $\{f_{|E} < a\} = A \bigcap E$. Now as $E \in \mathcal{A}$ we have $A \bigcap E \in \mathcal{A}$ so that $\{f_{|E} < a\} \in \mathcal{A}$

as $f_{|E}$ is $\mathcal{A}$-measurable we have by 18.98 that $\{a \leqslant f_{|E}\} \in \mathcal{A}_{|\operatorname{dom}(f_{|E})} = \mathcal{A}_{|E}$ so there exists a $A \in \mathcal{A}$ such that $\{a \leqslant f_{|E}\} = A \bigcap E$. Now as $E \in \mathcal{A}$ we have $A \bigcap E \in \mathcal{A}$ so that $\{a \leqslant f_{|E}\} \in \mathcal{A}$. □

The following lemma gives a simpler condition to check if a function is a positive integrable function.

Lemma 18.130. *Let $\langle X, \mathcal{A}, \mu \rangle$ be a measure space and $f: X \rightarrow \mathbb{R}$ a partial function such that $f(X) \subseteq [0, \infty[$ then we have that*

1. *$f \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ if and only if there exists a conegligible set $E \in \mathcal{A}$ such that*
   a. *$E \subseteq \operatorname{dom}(f)$*
   b. *$f_{|E}$ is $\mathcal{A}$-measurable*

c. *$\forall \varepsilon > 0$ $\mu(\{\varepsilon \leqslant f_{|E}\}) < \infty$ [well defined as $\{\varepsilon \leqslant f_{|E}\} \in \mathcal{A}$ because $E \in \mathcal{A}$ by (a), (b) combined with 18.129]*

d. $\sup\left(\left\{\int^S g d\mu : g \in \mathcal{S}[X, \mathcal{A}] \text{with } g \leqslant_{a.e.} f\right\}\right) < \infty$

2. *Let $f \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ and let $h: X \to \mathbb{R}$ be a partial function such that*

   a. $h(X) \subseteq [0, \infty[$

   b. *$\exists F \subseteq X$ conegligible such that $h_{|F}$ is $\mathcal{A}$-measurable*

   c. $h \leqslant_{a.e.} f$

   then $h \in \mathcal{L}_+[X, \mathcal{A}, \mu]$

Proof.

1. Let $f \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ then there exists a sequence $\{f_n\}_{n \in \mathbb{N}}$ of simple functions such that $\forall n \in \mathbb{N}$ $0 \leqslant f_n \leqslant f_{n+1}$, $\sup\left(\left\{\int^S f_n d\mu | n \in \mathbb{N}\right\}\right) < \infty$ and $\lim_{n \to \infty} f_n(x) =_{a.e.} f(x)$. Then $\left\{x \in \operatorname{dom}(f) | \lim_{n \to \infty} f_n(x) = f(x)\right\} \subseteq \operatorname{dom}(f)$ is conegligible and using 18.27(4) there exists then a conegligible set $E \in \mathcal{A}$ such that $E \subseteq \left\{x \in \operatorname{dom}(f) | \lim_{n0 < \infty} f_n(x) = f(x)\right\} \subseteq \operatorname{dom}(f)$. Now $\forall x \in E$ we have that $f_{|E}(x) = f(x) = \lim_{n \to \infty} f_n(x)$ hence $f_{|E} = \left(\lim_{n \to \infty} f_n\right)_{|E}$ which is $\mathcal{A}$-measurable [because $\forall n \in \mathbb{N}$ f_n is $\mathcal{A}$-measurable (see 18.115) so that (see 18.104) $\lim_{n \to \infty} f_n$ is $\mathcal{A}$-measurable and finally using 18.103(7) $\left(\lim_{n \to \infty} f_n\right)_{|E}$ is $\mathcal{A}$-measurable]. To summarize

   $$\exists E \in \mathcal{A},\ E \subseteq \operatorname{dom}(f),\ E \text{ is conegligible and } f_{|E} = \lim_{n \to \infty} f_n \text{ is } \mathcal{A}\text{-measurable} \tag{18.237}$$

   This proves (1.a.) and (1,b.). Next take $\varepsilon > 0$ and set $H_n = \left\{\frac{1}{2} \cdot \varepsilon \leqslant f_n\right\} \bigcap E$. As f_n is a simple function thus measurable [see 18.115] proving that $\left\{\frac{1}{2} \cdot \varepsilon \leqslant f_n\right\} \in \mathcal{A}$ which together with $E \in \mathcal{A}$ gives $H_n \in \mathcal{A}$. Define the simple function $h_n = \left(\frac{1}{2} \cdot \varepsilon\right) \cdot \mathcal{X}_{H_n}$, then $\forall x \in X$ we have either

   $\boldsymbol{x \in H_n}$**.** then $\frac{1}{2} \cdot \varepsilon \leqslant f(x) \underset{H_n \subseteq E}{=} f_{n|E}(x)$ so that $h_n(x) = \frac{1}{2} \cdot \varepsilon \cdot \mathcal{X}_{H_n}(x) = \frac{1}{2} \cdot \varepsilon \leqslant f_n(x)$

   $\boldsymbol{x \in H_n}$**.** then $h_n(x) = 0 \leqslant f_n(x)$ [as $0 \leqslant f_n$]

   proving that $\frac{1}{2} \cdot \varepsilon \cdot \mathcal{X}_{H_n} \leqslant f_n$. Using 18.120 we have then that $\frac{1}{2} \cdot \varepsilon \cdot \mu(H_n) = \int^S \left(\frac{1}{2} \cdot \varepsilon \cdot \mathcal{X}_{H_n}\right) d\mu \leqslant \int^S f_n d\mu$. Hence

   $$\forall n \in \mathbb{N} \text{ we have } \frac{1}{2} \cdot \varepsilon \cdot \mu(H_n) \leqslant \int^S f_n d\mu \leqslant \sup\left(\left\{\int^S f_n d\mu | n \in \mathbb{N}\right\}\right) < \infty \tag{18.238}$$

   or taking the supremum

   $$\sup\left(\{\mu(H_n) | n \in \mathbb{N}\}\right) \leqslant \frac{2}{\varepsilon} \cdot \sup\left(\left\{\int^S f_n d\mu | n \in \mathbb{N}\right\}\right) < \infty \tag{18.239}$$

Take $n \in \mathbb{N}$ then if $x \in H_n$ we have $\frac{1}{2} \cdot \varepsilon \leqslant f_n(x) \leqslant f_{n+1}(x) \Rightarrow x \in H_{n+1}$ proving that $\{H_n\}_{n \in \mathbb{N}}$ is increasing. Using 18.20 it follows that $\mu(\bigcup_{n \in \mathbb{N}} H_n) = \sup(\{\mu(H_n) | n \in \mathbb{N}\})$ which if we combine this with (18.239) gives

$$\mu\left(\bigcup_{n \in \mathbb{N}} H_n\right) \leqslant \frac{2}{\varepsilon} \cdot \sup\left(\left\{\int^S f_n d\mu | n \in \mathbb{N}\right\}\right) < \infty \tag{18.240}$$

If now $x \in \{\varepsilon \leqslant f_{|E}\}$ then $x \in E$ and $\frac{\varepsilon}{2} < \varepsilon \leqslant f_{|E}(x) = \lim_{n \to \infty} f_n(x) \underset{\{f_n(x)\}_{n \in \mathbb{N}} \text{ increasing} + 12.354}{=} \sup(\{f_n(x) | n \in \mathbb{N}\})$ so there exists a $n \in \mathbb{N}$ such that $\frac{\varepsilon}{2} < f_n(x) \leqslant f_{|E}(x) \Rightarrow x \in \left\{\frac{1}{2} \cdot \varepsilon \leqslant f_n\right\} \bigcap E = H_n$ which proves that $\{\varepsilon \leqslant f_{|E}\} \subseteq \bigcup_{n \in \mathbb{N}} H_n$ hence as also $\{\varepsilon \leqslant f_{|E}\} \in \mathcal{A}$ [see (18.237) and 18.129] we have

$$\mu(\{\varepsilon \leqslant f_{|E}\}) \leqslant \mu\left(\bigcup_{n \in \mathbb{N}} H_n\right) < \infty \tag{18.241}$$

which proves (1.c). Using 18.125 we have that $\sup\left(\left\{\int^S g\, d\mu | g \in \mathcal{S}[X, \mathcal{A}] \text{ such that } g \leqslant_{a.e.} f\right\}\right) = \sup\left(\left\{\int^S f_n d\mu | n \in \mathbb{N}\right\}\right) < \infty$ which proves 1.d. Hence we have proved that if $f \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ that there exists a $E \in \mathcal{A}$ such that (1.a),(1.b),(1.c) and (1.d) is satisfied.

For the opposite implication, assume that there exists a $E \in \mathcal{A}$ so that $(1.a)$, $(1.b)$, $(1.c)$ and $(1.d)$ are satisfied. Using the previous lemma [see 18.128] on $f_{|E}$ there exists a sequence of functions $\{f_n\}_{n \in \mathbb{N}}$ defined as

$$\forall n \in \mathbb{N}\, f_n: X \to \mathbb{R} \text{ is defined as } \sum_{k=1}^{4^n} \frac{1}{2^n} \cdot \mathcal{X}_{\left\{\frac{k}{2^n} \leqslant f_{|E}\right\}} \tag{18.242}$$

such that

$$\forall n \in \mathbb{N} \text{ we have } 0 \leqslant f_n \leqslant f_{n+1} \text{ and } f_{|E} = \lim_{n \to \infty} f_n = \sup(\{f_n | n \in \mathbb{N}\}) \tag{18.243}$$

As by (1.a),(1.b) $f_{|E}$ is measurable and $E \in \mathcal{A}$ it follows from 18.129 that $\left\{\frac{k}{2^n} \leqslant f_{|E}\right\} \in \mathcal{A}$, further using (1.c) $\mu\left(\left\{\frac{k}{2^n} \leqslant f_{|E}\right\}\right) < \infty$. So we have that $\{f_n\}_{n \in \mathbb{N}} \subseteq \mathcal{S}[X, \mathcal{A}]$. Further as E is conegligible by assumption we have that

$$\exists \{f_n\}_{n \in \mathbb{N}} \subseteq \mathcal{S}[X, \mathcal{A}] \text{ with } 0 \leqslant f_n \leqslant f_{n+1} \text{and } f =_{a.e.} \lim_{n \to \infty} f_n \tag{18.244}$$

Now $\forall n \in \mathbb{N}$ we have that $\forall x \in E\; f_n(x) \leqslant \sup(\{f_n(x) | n \in \mathbb{N}\}) \underset{(18.243)}{=} f_{|E}(x) = f(x)$ which as E is conegligible means that $f_n \leqslant_{a.e.} f$ so that $\int^S f_n d\mu \in \left\{\int^S g d\mu | g \in S[X, \mathcal{A}] \wedge g \leqslant_{a.e} f\right\}$. Hence

$$\sup\left(\left\{\int^S f_n d\mu | n \in \mathbb{N}\right\} \leqslant \sup\left(\left\{\int^S g\; d\mu | g \in S[X, \mathcal{A}] \wedge g \leqslant_{a.e} f\right\}\right) <_{(1.d)} \infty \tag{18.245}$$

Using then(18.244),(18.245) and the fact that $f(X)\subseteq[0,\infty[$ and $\operatorname{dom}(f)$ is conegligible we have that

$$f\in\mathcal{L}_+[X,\mathcal{A},\mu] \tag{18.246}$$

which proves the reverse implication.

2. As $f\in\mathcal{L}_+[X,\mathcal{A},\mu]$ there exists by (1.a),(1.b),(1,c),(1,d) a conegligible set $E\in\mathcal{A}$ with $E\subseteq\operatorname{dom}(f)$, $f_{|E}$ $\mathcal{A}$-measurable, $\mu(\{\varepsilon\leqslant f_{|E}\})\forall\varepsilon>0$ and $\sup\left(\left\{\int^S gd\mu: g\in\mathcal{S}[X,\mathcal{A}]\text{with }g\leqslant_{a.e.}f\right\}\right)<\infty$. Further $\{x\in\operatorname{dom}(h)\bigcap\operatorname{dom}(f)|h(x)\leqslant f(x)\}$ is conegligible [as $g\leqslant_{a.e.}f$] and F is conegligible so that $E\bigcap F\bigcap\{x\in\operatorname{dom}(h)\bigcap\operatorname{dom}(f)|h(x)\leqslant f(x)\}$ is conegligible. So using 18.27 there exist a conegligible set $E'\in\mathcal{A}$ such that $E'\subseteq E\bigcap F\bigcap\{x\in\operatorname{dom}(h)\bigcap\operatorname{dom}(f)|h(x)\leqslant f(x)\}\subseteq\operatorname{dom}(f),\operatorname{dom}(h)$. Now as the partial function $h_{|F}$ is $\mathcal{A}$-measurable and $h_{|E'}=(h_{|F})_{|E'}$ it follows from 18.103 (7) that $h_{|E'}$ is $\mathcal{A}$-measurable. To summarize

$$\exists E'\in\mathcal{A}\text{ with }E'\subseteq\operatorname{dom}(h)\wedge h_{|E'}\text{ is }\mathcal{A}\text{-measurable} \tag{18.247}$$

Take $\varepsilon>0$ then if $x\in\{\varepsilon<h_{|E'}\}$ we have $x\in\operatorname{dom}(h)\bigcap E'\wedge\varepsilon\leqslant h_{|E'}(x)=h(x)$ then as $E'\subseteq\{x\in\operatorname{dom}(h)\bigcap\operatorname{dom}(f)|h(x)\leqslant f(x)\}\bigcap E$ we have that $x\in\operatorname{dom}(f)\bigcap E\wedge\varepsilon\leqslant f(x)=f_{|E}(x)$ so that $x\in\{\varepsilon\leqslant f_{|E}\}$. Hence $\{\varepsilon\leqslant h_{|E'}\}\subseteq\{\varepsilon\leqslant f_{|E}\}$ giving by 18.20 that

$$\forall\varepsilon>0\text{ we have }\mu(\{\varepsilon\leqslant h_{|E'}\})\leqslant\mu(\{\varepsilon\leqslant f_{|E}\})<\infty \tag{18.248}$$

If $g\in\mathcal{S}[X,\mathcal{A}]$ such that $g\leqslant_{a.e.}h$ then as $h\leqslant_{a.e.}f$ we have by 18.30 that $g\leqslant_{a.e}f$ from which it follows that $\left\{\int^S gd\mu|g\in\mathcal{S}[X,\mathcal{A}]\wedge g\leqslant_{a.e.}h\right\}\subseteq\left\{\int^S gd\mu|g\in\mathcal{S}[X,\mathcal{A}]\wedge g\leqslant_{a.e.}f\right\}$. Applying then 2.171 gives

$$\sup\left(\left\{\int^S gd\mu|g\in\mathcal{S}[X,\mathcal{A}]\wedge g\leqslant_{a.e.}h\right\}\right)\leqslant\sup\left(\left\{\int^S gd\mu|g\in\mathcal{S}[X,\mathcal{A}]\wedge g\leqslant_{a.e.}f\right\}\right)<\infty \tag{18.249}$$

Finally (18.247),(18.248) and (18.249) are equivalent with (1.a),(1.b),(1.c) and (1.d) so that

$$h\in\mathcal{L}_+[X,\mathcal{A},\mu]$$

□

$\mathcal{L}_+[X,\mathcal{A},\mu]$ is called the set of positive integrable functions so it make sense to define the integral for $f\in\mathcal{L}_+[X,\mathcal{A},\mu]$. This is done in the following definition.

Definition 18.131. Let $\langle X,\mathcal{A},\mu\rangle$ *be a measure space,* $f\in\mathcal{L}_+[X,\mathcal{A},\mu]$ *then the integral of* f *noted by* $\int^+ fd\mu$ *is defined as*

$$\int^+ fd\mu=\sup\left(\left\{\int^S gd\mu|g\in\mathcal{S}[X,\mathcal{A}]\;\textit{ and }\;g\leqslant_{a.e.}f\right\}\right)<\infty\text{ [see }\mathit{18.130}]$$

Proposition 18.132. *Let* $f\in\mathcal{L}_+[X,\mathcal{A},\mu]$ *then* $\forall\{f_i\}_{i\in\mathbb{N}}\subseteq\mathcal{S}[X,\mathcal{A}]$ *such that* $\forall n\in\mathbb{N}\ 0\leqslant f_n\leqslant f_{n+1}$ *[a increasing sequence of non negative simple functions] and* $f=_{a.e.}\lim_{n\to\infty}f_n$ *we have* $\int^+ fd\mu=\lim_{n\to\infty}\int^S f_nd\mu=\sup\left(\left\{\int^S f_nd\mu|n\in\mathbb{N}\right\}\right)$

Proof. As $f =_{a.e.} \lim_{n\rightarrow\infty} f_n$ we have that $E = \left\{x \in \mathrm{dom}(f) | f(x) = \lim_{n\rightarrow\infty} f_n(x)\right\}$ is conegligible. Hence if $x \in E$ then $f(x) = \lim_{n\rightarrow\infty} f_n(x) \underset{\{f_n(x)\}_{n\in\mathbb{N}} \text{is increasing}}{=} \sup(\{f_n(x) | n \in \mathbb{N}\})$ so that $\forall x \in E$ we have that $f_n(x) \leqslant \sup(\{f_n(x)|n \in \mathbb{N}\}) = f(x)$ proving that $\forall n \in \mathbb{N}$ we have that $f_n \leqslant_{a.e.} f$. So $\left\{\int^S f_n d\mu | n \in \mathbb{N}\right\} \subseteq \left\{\int^S g\, d\mu | g \in \mathcal{S}[X, \mathcal{A}] \wedge g \leqslant_{a.e.} f\right\}$ and thus $\sup\left(\left\{\int^S f_n d\mu | n \in \mathbb{N}\right\}\right) \leqslant \sup\left(\left\{\int^S g d\mu | g \in \mathcal{S}[X, \mathcal{A}] \wedge g \leqslant_{a.e} f\right\}\right)$. As by 18.130 (d) $\sup\left(\left\{\int^S g d\mu | g \in \mathcal{S}[X, \mathcal{A}] \wedge g \leqslant_{a.e} f\right\}\right) < \infty$ we conclude that

$$\sup\left(\left\{\int^S f_n d\mu | n \in \infty\right\}\right) < \infty$$

this together with the rest of the conditions on $\{f_n\}_{n\in\mathbb{N}}$ assumed in this lemma allows us to apply 18.125 giving that

$$\sup\left(\left\{\int^S f_n d\mu | n \in \infty\right\}\right) = \sup\left(\left\{\int^S g d\mu | g \in \mathcal{S}[X, \mathcal{A}] \wedge g \leqslant_{a.e} f\right\}\right) \underset{\text{def}}{=} \int^+ f d\mu \quad \square$$

We prove now that $\int^+$ is a extension for $\int^S$

Proposition 18.133. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space and $f \in \mathcal{S}[X, \mathcal{A}]$ such that $f(X) \subseteq [0, \infty[$ then $f \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ and $\int^+ f d\mu = \int^S f d\mu$*

Proof. Define $\{f_n\}_{n\in\mathbb{N}}$ by $f_n = f$ then and we have $\forall n \in \mathbb{N}$ we have $0 \leqslant f_n \leqslant f_{n+1}$ and $\sup(\{\int f_n d\mu | n \in \mathbb{N}\}) = \sup(\{\int f d\mu\}) = \int f d\mu < \infty$. Further $\forall x \in X$ $\lim_{n\rightarrow\infty} f_n(x) = f(x)$ which as X is conegligible proves that $\lim_{n\rightarrow\infty} f_n =_{a.e.} f_\varepsilon$. So by definition of $\mathcal{L}_+[X, \mathcal{A}, \mu]$ it follows that

$$f \in \mathcal{L}_+[X, \mathcal{A}, \mu]$$

Finally

$$\int^+ f d\mu \underset{18.132}{=} \lim_{n\rightarrow\infty} \int^S f_n d\mu = \int^S f_n d\mu \qquad \square$$

Corollary 18.134. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space then the function $C_0: X \rightarrow \mathbb{R}$ defined by $C_0(x) = 0$ is a element of $\mathcal{L}$ and $\int C_0 d\mu = 0$*

Proof. This follows from 18.133 and the fact that $C_0 = \mathcal{X}_\emptyset$ is a simple function such that $\int^S C_0 d\mu = 0$ $\square$

Lemma 18.135. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space then*

1. *$\forall f, g \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ we have that $f + g \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ and $\int^+ (f + g) d\mu = \int^+ f d\mu + \int^+ g d\mu$*
2. *$\forall f \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ and $c \in [0, ..., \infty[$ that $\int^+ (c \cdot f) d\mu = c \cdot \int^+ f d\mu$*
3. *If $f, g \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ and $f \leqslant_{a.e.} g$ then $\int^+ f d\mu \leqslant \int^+ g d\mu$*
4. *If $f \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ and $g: X \rightarrow \mathbb{R}$ a partial function, $g(X) \subseteq [0, \infty[$ and $f =_{a.e.} g$ then $g \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ and $\int^+ f d\mu = \int^+ g d\mu$*
5. *If f_1, g_1, f_2, $g_2 \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ such that $f_1 - g_1 = f_2 - g_2$ then $\int^+ f_1 d\mu - \int^+ g_1 d\mu = \int^+ f_2 d\mu - \int^+ g_2 d\mu$*

Proof.

1. As $f, g \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ we have $f(X) \subseteq [0, \infty[, g(X) \subseteq [0, \infty[, \mathrm{dom}(f), \mathrm{dom}(g)$ are conegligible and there exists $\{f_n\}_{n\in\mathbb{N}}, \{g_n\}_{n\in\mathbb{N}} \subseteq \mathcal{S}[X, \mathcal{A}]$ with $\forall n \in \mathbb{N}\ 0 \leqslant f_n \leqslant f_{n+1}, 0 \leqslant g_n \leqslant g_{n+1}, f =_{a.e.} \lim_{n\rightarrow\infty} f_n$, $g =_{a.e.} \lim_{n\rightarrow\infty} g_n$ and $\sup\left(\left\{\int^S f_n d\mu | n \in \mathbb{N}\right\}\right) < \infty$, $\sup\left(\left\{\int^S g_n d\mu\right\}\right) < \infty$. Then trivially we have

$$\forall n \in \mathbb{N} \text{ we have } 0 \leqslant f_n + g_n \leqslant f_{n+1} + g_{n+1} \tag{18.250}$$

Similar we have as $f =_{a.e.} \lim_{n\rightarrow\infty} f_n$ and $g =_{a.e.} \lim_{n\rightarrow\infty} g_n$ that $F = \left\{x \in \mathrm{dom}(f) \bigcap \mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right) | \lim_{n\rightarrow\infty} f_n(x) = f(x)\right\}$, $G = \left\{x \in \mathrm{dom}(g) \bigcap \mathrm{dom}\left(\lim_{n\rightarrow\infty} g_n\right) | \lim_{n\rightarrow\infty} g_n(x) = g(x)\right\}$ are conegligible so that $F \bigcap G$ is conegligible. As $\forall x \in F \bigcap G$ we have $\lim_{n\rightarrow\infty} f_n(x) + \lim_{n\rightarrow\infty} g_n(x) = \lim_{n\rightarrow\infty} (f_n(x) + g_n(x)) = \lim_{n\rightarrow\infty} (f_n + g_n)(x)$ we have that

$$\lim_{n\rightarrow\infty} (f_n + g_n) =_{a.e.} f + g \tag{18.251}$$

Next

$$\begin{aligned}
\sup\left(\left\{\int^S (f_n + g_n) | n \in \mathbb{N}\right\}\right) &\underset{18.123}{=} \lim_{n\rightarrow\infty} \int^S (f + g) d\mu \\
&\underset{18.120}{=} \lim_{n\rightarrow\infty} \left(\int^S f_n d\mu + \int^S g_n d\mu\right) \\
&\underset{12.341}{=} \lim_{n\rightarrow\infty} \int^S f_n d\mu + \lim_{n\rightarrow\infty} \int^S g_n d\mu \\
&\underset{18.132}{=} \int^+ f d\mu + \int^+ g d\mu & (18.252) \\
&< \infty & (18.253)
\end{aligned}$$

Using (18.250),,(18.251),(18.253) we have by the definition of $\mathcal{L}_+[X, \mathcal{A}, \mu]$ that

$$f + g \in \mathcal{L}_+[X, \mathcal{A}, \mu] \tag{18.254}$$

Finally using 18.132 we have that $\sup\left(\left\{\int^S (f_n + g_n) | n \in \mathbb{N}\right\}\right) = \int^+ (f + g) d\mu$ so using (18.252) we have

$$\int^+ (f + g) d\mu = \int^+ f d\mu + \int^+ g d\mu$$

2. As $f \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ we have that $f(X) \subseteq [0, \infty[$ and there exists $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathcal{S}[X, \mathcal{A}]$ with $\forall n \in \mathbb{N}\ 0 \leqslant f_n \leqslant f_{n+1}$, $f =_{a.e.} \lim_{n\rightarrow\infty} f_n$ and $\sup\left(\left\{\int^S f_n d\mu | n \in \mathbb{N}\right\}\right) < \infty$. As $c \geqslant 0$ we have that

$$\forall n \in \mathbb{N} \text{ we have } 0 \leqslant c \cdot f_n \leqslant c \cdot f_{n+1} \tag{18.255}$$

Further $\forall x \in \left\{x \in \operatorname{dom}(f) \bigcap \operatorname{dom}\left(\lim_{n \rightarrow \infty} f_n\right) | \lim_{n \rightarrow \infty} f_n(x) = f(x)\right\}$ a conegligible set we have that $\lim_{n \rightarrow \infty} (c \cdot f_n(x)) \underset{12.341}{=} c \cdot \lim_{n \rightarrow \infty} f_n(x)$ so that

$$\lim_{n \rightarrow \infty} f_n =_{a.e.} f \tag{18.256}$$

Further

$$\begin{aligned}
\sup\left(\left\{\int^S (c \cdot f_n) d\mu | n \in \mathbb{N}\right\}\right) &\underset{18.123}{=} \lim_{n \rightarrow \infty}\left(\int^S (c \cdot f_n) d\mu\right) \\
&= \lim_{n \rightarrow \infty}\left(c \cdot \int^S f_n d\mu\right) \\
&\underset{12.341}{=} c \cdot \lim_{n \rightarrow \infty} \int^C f_n d\mu \\
&\underset{18.132}{=} c \cdot \int^+ f_n d\mu \qquad (18.257) \\
&< \infty \qquad (18.258)
\end{aligned}$$

Using (18.255),(18.256),(18.258) we have by the definition of $\mathcal{L}_+[X, \mathcal{A}, \mu]$ that

$$c \cdot f \in \mathcal{L}_+[X, \mathcal{A}, \mu] \tag{18.259}$$

Finally using 18.132 we have that $\sup\left(\left\{\int^S (c \cdot f_n) | n \in \mathbb{N}\right\}\right) = \int^+ (c \cdot f) d\mu$ so using (18.257) we have

$$\int^+ (c \cdot f) d\mu = c \cdot \int^+ f d\mu$$

3. As $f, g \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ with $f \leqslant_{a.e.} g$. Let now $h \in \mathcal{S}[X, \mathcal{A}]$ such that $h \leqslant_{a.e.} f$ then by 18.30 we have $h \leqslant_{a.e.} g$. Hence $\left\{\int^S h d\mu | h \in \mathcal{S}[X, \mathcal{A}] \wedge h \leqslant_{a.e.} f\right\} \subseteq \left\{\int^S h d\mu | h \in \mathcal{S}[X, \mathcal{A}] \wedge h \leqslant_{a.e} g\right\}$ so that

   $$\int^+ f d\mu = \sup\left(\left\{\int^S h d\mu | h \in \mathcal{S}[X, \mathcal{A}] \wedge h \leqslant_{a.e.} f\right\}\right) \leqslant \sup\left(\left\{\int^S h d\mu | h \in \mathcal{S}[X, \mathcal{A}] \wedge h \leqslant_{a.e.} g\right\}\right) = \int^+ g d\mu$$

4. As $f \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ we have that $f(X) \subseteq [0, \infty[, \operatorname{dom}(f)$ and there exists $\{f_n\}_{n \in \mathbb{N}} \subseteq \mathcal{S}[X, \mathcal{A}]$ with $\forall n \in \mathbb{N}\ \ 0 \leqslant f_n \leqslant f_{n+1}$, $f =_{a.e.} \lim_{n \rightarrow \infty} f_n$ and $\sup\left(\left\{\int^S f_n d\mu | n \in \mathbb{N}\right\}\right) < \infty$. Given that $f =_{a.e.} g$ we have by 18.30 that $\lim_{n \rightarrow \infty} f_n =_{a.e.} g$. So by definition we have

   $$g \in \mathcal{L}_+[X, \mathcal{A}, \mu]$$

   Further we have that

   $$\int^+ f d\mu \underset{18.132}{=} \lim_{n \rightarrow \infty} \int^S f_n d\mu \underset{18.132}{=} \int^+ g d\mu$$

5. As f_1, g_1, f_2, g_2 we have that $f_1 + g_2, f_2 + g_1 \in \mathcal{L}_1$. Further $\forall x \in \mathrm{dom}(f_1) \bigcap \mathrm{dom}(f_2) \bigcap \mathrm{dom}(g_1) \bigcap \mathrm{dom}(g_2)$ [a conegligible set (see 18.127)] we have $(f_1 + g_2)(x) = f_1(x) + g_2(x) = f_2(x) + g_1(x) = (f_2 + g_1)(x)$ proving that $f_1 + g_2 =_{a.e.} f_2 + g_1$ so that $\int^+ f_1 d\mu + \int^+ g_2 d\mu \underset{(1)}{=} \int^+ (f_1 + g_2) d\mu \underset{(4)}{=} \int^+ (f_2 + g_1) d\mu = \int^+ f_2 d\mu + \int^+ g_1 d\mu$ from which it follows that

$$\int^+ f_1 d\mu - \int^+ g_1 d\mu = \int^+ f_2 d\mu - \int^+ g_2 d\mu$$

□

18.2.2.3 Measure integral

We extend now the definition of real functions that do not have to be positive.

Definition 18.136. *Let $\langle X, \mathcal{A}, \mu \rangle$ be a measure space then a partial function $f: X \to \mathbb{R}$ is **integrable** if there exists $f_1, f_2 \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ such that $f = f_1 - f_2$ [here we assume that $\mathrm{dom}(f) = \mathrm{dom}(f_1) \bigcap \mathrm{dom}(f_2)$] The **integral** of a **integrable** function is defined as $\int f d\mu = \int^+ f_1 d\mu - \int^+ f_2 d\mu$ [this is well defined because of 18.135]. The set of integrable functions is noted as $\mathcal{L}[X, \mathcal{A}, \mu]$*

Remark 18.137. As $\mathrm{dom}(f) = \mathrm{dom}(f_1) \bigcap \mathrm{dom}(f_2)$ and $f_1, f_2 \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ so that by 18.127 $\mathrm{dom}(f_1), \mathrm{dom}(f_2)$ are conegligible it follows that $\mathrm{dom}(f)$ is negligible.

Note 18.138. If $f \in \mathcal{L}[\mathbb{R}, \mathcal{L}, \lambda]$ [or $f \in \mathcal{L}[\mathbb{R}^n, \mathcal{L}^n, \lambda^n]$ [see 18.72 and Lebesgue measurable set] then f is is said to be Lebesgue integrable and $\int f d\lambda$ [or $\int f d\lambda^n$] is the Lebesgue integral of f.

Notation 18.139. *A other notation that sometimes will be used is $\int f(x) d\mu(x)$ which will be useful in the following cases:*

1. *To avoid excessive notation. Instead of saying: Take $\int f d\mu$ where f is the function $f: \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x \cdot \cos(x)$ we can just say take $\int x \cdot \cos(x) d\mu(x)$*
2. *Parameterized functions. Instead of saying: Let $f: X \times Y \to \mathbb{R}$ be a partial function, given $x \in X$ define $g_x: Y \to \mathbb{R}$ by $g_x(y) = f(x, y)$ take $\int g_x d\mu$ we can just say take $\int f(x, y) d\mu(y)$*

of course there is some ambiguity in this notation as $\int f(x) d\mu(x)$ is the same as $\int f(y) d\mu(y)$ so we will use this notation only in the cases (1) and (2).

We show now that $\int$ extends $\int^+$ and thus also $\int^S$

Proposition 18.140. *Let $\langle X, \mathcal{A}, \mu \rangle$ be a measure space then we have*

1. *If $f \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ then $f \in \mathcal{L}[X, \mathcal{A}, \mu]$ and $\int f d\mu = \int^+ f d\mu$*
2. *If $f \in \mathcal{S}[X, \mathcal{A}]$ then $f \in \mathcal{L}[X, \mathcal{A}, \mu]$ and $\int f d\mu = \int^S f d\mu$*

Proof.

1. If $f \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ then $f = f - C_0$ where $f, C_0 \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ and $\operatorname{dom}(f) = \operatorname{dom}(f) \bigcap X = \operatorname{dom}(f) \bigcap \operatorname{dom}(C_0)$ hence $\int f d\mu = \int^+ f d\mu - \int^+ C_9 d\mu \underset{18.134}{=} \int^+ f d\mu$

2. Let $f \in \mathcal{S}[X, \mathcal{A}]$ then there exists a representation $f = \sum_{i=1}^n \alpha_i \cdot \mathcal{X}_{A_i}$. Define then $\beta_i = \max(\alpha_i, 0) \geqslant 0$ and $\gamma_i = -\min(0, \alpha_i) \geqslant 0$ then for $\beta_i - \gamma_i$ we have the following cases to consider

   $\boldsymbol{0 \leqslant \alpha_i}$. then $\beta_i - \gamma_i = \max(\alpha_i, 0) + \min(0, \alpha) = \alpha_i + 0 = \alpha_i$

   $\boldsymbol{\alpha_i < 0}$. then $\beta_i - \gamma_i = \max(\alpha_i, 0) + \min(0, \alpha) = 0 + \alpha_i = \alpha_i$

   Hence

$$
\begin{aligned}
f = \sum_{i=1}^n \alpha_i \cdot \mathcal{X}_{A_i} &= \sum_{i=1}^n (\beta_i - \gamma_i) \cdot \mathcal{X}_{A_i} \\
&= \sum_{i=1}^n \beta_i \cdot \mathcal{X}_{A_i} - \sum_{i=1}^n \gamma_i \cdot \mathcal{X}_{A_i} \\
&= f_1 - f_2
\end{aligned}
$$

   where $f_1 = \sum_{i=1}^n \beta_i \cdot \mathcal{X}_{A_i}$, $f_2 = \sum_{i=1}^n \gamma_i \cdot \mathcal{X}_{A_i}$ are elements of $\mathcal{L}_+[X, \mathcal{A}, \mu]$ [using 18.133] so that $f \in \mathcal{L}[X, \mathcal{A}, \mu]$. Further

$$
\begin{aligned}
\int f d\mu &= \int^+ f_1 d\mu - \int^+ f_2 \, d\mu \\
&\underset{18.133}{=} \int^S f_1 d\mu - \int^S f_2 d\mu \\
&= \sum_{i=1}^n \beta_i \cdot \mu(A_i) - \sum_{i=1}^n \gamma_i \cdot \mu(A_i) \\
&= \sum_{i=1}^n (\beta_i - \gamma_i) \cdot \mu(A_i) \\
&= \sum_{i=1}^n \alpha_i \cdot \mu(A_i) \\
&= \int^S f d\mu
\end{aligned}
$$

□

Theorem 18.141. *Let $\langle X, \mathcal{A}, \mu \rangle$ be a measure space then*

1. *$\forall f, g \in \mathcal{L}[X, \mathcal{A}, \mu]$ we have that $f + g \in \mathcal{L}[X, \mathcal{A}, \mu]$ and $\int (f + g) d\mu = \int f d\mu + \int g d\mu$*
2. *$\forall f \in \mathcal{L}[X, \mathcal{A}, \mu]$ and $c \in \mathbb{R}$ we have that $c \cdot f \in \mathcal{L}[X, \mathcal{A}, \mu]$ and $\int (c \cdot f) d\mu = c \cdot \int f d\mu$*
3. *$\forall f \in \mathcal{L}[X, \mathcal{A}, \mu]$ with $0 \leqslant_{a.e.} f$ we have $0 \leqslant \int f d\mu$*
4. *$\forall f \in \mathcal{L}[X, \mathcal{A}, \mu]$ with $f \leqslant_{a.e.} 0$ we have $\int f d\mu \leqslant 0$*

5. *$\forall f, g \in \mathcal{L}[X, \mathcal{A}, \mu]$ with $f \leqslant_{a,e,} g$ we have $\int f d\mu \leqslant \int g d\mu$*

Proof.

1. As $f, g \in \mathcal{L}[X, \mathcal{A}, \mu]$ there exists $f_1, f_2, g_1, g_2 \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ such that $f = f_1 - f_2, g = g_1 - g_2, \int f d\mu = \int^+ f_1 d\mu - \int^+ f_2 d\mu$ and $\int g d\mu = \int^+ g_1 d\mu - \int^+ g_2 d\mu$. Using 18.135 we have that $f_1 + f_2, g_1 + g_2 \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ and $\int (f_1 + f_2) d\mu = \int^+ f_1 d\mu + \int^+ f_2 d\mu, \int^+ (g_1 + g_2) d\mu = \int^+ g_1 d\mu + \int^+ g_2 d\mu$. Finally $f + g = (f_1 + f_2) - (g_1 + g_2)$ proving that $f + g \in \mathcal{L}[X, \mathcal{A}, \mu]$ and $\int (f + g) d\mu = \int^+ (f_1 + f_2) d\mu - \int^+ (g_1 + g_2) d\mu = \int^+ f_1 d\mu - \int^+ f_2 d\mu + \int^+ g_1 d\mu - \int^+ g_2 d\mu = \int f d\mu + \int g d\mu$
2. As $f \in \mathcal{L}[X, \mathcal{A}, \mu]$ there exists $f_1, f_2 \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ such that $f = f_1 - f_2$ and $\int f d\mu = \int^+ f_1 d\mu - \int^+ f_2 d\mu$. Now $-f = -(f_1 - f_2) = f_2 - f_1$ so that

$$-f \in \mathcal{L}[X, \mathcal{A}, \mu] \text{ and } \int (-f) d\mu = \int^+ f_2 d\mu - \int^+ f_1 d\mu = -\int f d\mu \tag{18.260}$$

For c we consider now two cases

$0 \leqslant c$. then using 18.135 $c \cdot f_1, c \cdot f_2 \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ and $\int^+ (c \cdot f_1) d\mu$, $\int^+ (c \cdot f_2) d\mu$. As $c \cdot f = c \cdot f_1 - c \cdot f_2$ we have that $f \cdot c \in \mathcal{L}[X, \mathcal{A}, \mu]$ and $\int (c \cdot f) d\mu = \int^+ (c \cdot f_1) d\mu - \int^+ (c \cdot f_2) d\mu = c \cdot \int^+ f_1 d\mu - c \cdot \int^+ f_2 d\mu = c \cdot (\int^+ f_1 d\mu - \int^+ f_2 d\mu) = c \cdot \int f d\mu$

$c < 0$. then $0 < -c$ so using the previous case we have that $(-c \cdot f) \in \mathcal{L}[X, \mathcal{A}, \mu]$ and $\int (-c \cdot f) = (-c) \cdot \int f d\mu$. Using (18.260) we have that $c \cdot f = -(-c \cdot f) \in \mathcal{L}[X, \mathcal{A}]$ and $\int (c \cdot f) d\mu = \int -(-c \cdot f) d\mu = -\int (-c \cdot f) d\mu = -(-c) \cdot \int f d\mu = c \cdot \int f d\mu$

3. As $f \leqslant_{a.e.} 0$ we have that $\{x \in \operatorname{dom}(f) | f(x) \leqslant 0\}$ is μ-conegligible. Now as $f(x) \leqslant 0 \Leftrightarrow 0 \leqslant (-f(x))$ we have that $\{x \in \operatorname{dom}(f) | 0 \leqslant (-f)(x)\} = \{x \in \operatorname{dom}(f) | f(x) \leqslant 0\}$ proving that $0 \leqslant \int (-f) d\mu \underset{(2)}{=} -\int f d\mu$ from which it follows that $\int f d\mu \leqslant 0$.

$$x \in \{f\}$$

4. As $f \in \mathcal{L}[X, \mathcal{A}, \mu]$ there exists $f_1, f_2 \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ such that $f = f_1 - f_2$. As $0 \leqslant_{a.e.} f$ we have that $\{x \in \operatorname{dom}(f) | f_2 \leqslant f_1\} = \{x \in \operatorname{dom}(f) | 0 \leqslant f_2 - f_1\}$ is conegligible proving that $f_2 \leqslant_{a.e.} f_1$. Using 18.135 we have then that $\int f_2 d\mu \leqslant \int f_1 d\mu$ so that $\int f d\mu = \int f_1 d\mu - \int f_2 d\mu \geqslant 0$
5. As $f \leqslant_{a.e.} g$ we have that $\{x \in \operatorname{dom}(f) \bigcap \operatorname{dom}(g) | 0 \leqslant g(x) - f(x)\} = \{x \in \operatorname{dom}(f) \bigcap \operatorname{dom}(g) | f(x) \leqslant g(x)\}$ proving that $0 \leqslant_{a.e} g - f$ so by (3) we have $0 \leqslant \int (g - f) d\mu \underset{(1) \text{ and } (2)}{=} \int f d\mu - \int g d\mu$ proving that $\int f d\mu \leqslant \int g d\mu$

□

Lemma 18.142. *Let $f: X \to \mathbb{R}$ be a partial function then for the partial functions $f^+: X \to \mathbb{R}, f^-: X \to \mathbb{R}$ defined by $f^+ = \frac{1}{2} \cdot (|f| + f)$ and $f^- = \frac{1}{2} \cdot (|f| - f)$ with $\operatorname{dom}(f) = \operatorname{dom}(f^+) = \operatorname{dom}(f^-)$ we have*

1. *$f = f^+ - f^-$*

2. $|f| = f^+ + f^-$
3. $0 \leqslant f^+, f^-$

Proof.

1. $\forall x \in X$ we have $f^+(x) = \frac{1}{2} \cdot (|f(x)| + f(x)) - \frac{1}{2} \cdot (|f(x)| + f(x)) = f(x)$
2. $\forall x \in X$ we have $f^+(x) = \frac{1}{2} \cdot (|f(x)| + f(x)) + \frac{1}{2} \cdot (|f(x)| + f(x)) = |f(x)| = |f|(x)$
3. If $x \in X$ then we have either

   $0 \leqslant f(x)$. then

   a. $f^+(x) = \frac{1}{2} \cdot (|f(x)| + f(x)) = \frac{1}{2} \cdot (f(x) + f(x)) = f(x) \geqslant 0$

   b. $f^-(x) = \frac{1}{2} \cdot (|f(x)| - f(x)) = \frac{1}{2} \cdot (f(x) - f(x)) = 0 \geqslant 0$

   $f(x) < 0$. then

   a. $f^+(x) = \frac{1}{2} \cdot (|f(x)| + f(x)) = \frac{1}{2} \cdot (-f(x) + f(x)) = 0 \geqslant 0$

   b. $f^-(x) = \frac{1}{2} \cdot (|f(x)| - f(x)) = \frac{1}{2} \cdot (-f(x) - f(x)) = -f(x) \geqslant 0$ □

Theorem 18.143. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space and $f: X \to \mathbb{R}$ a partial function the the following are equivalent*

1. *f is integrable [or $f \in \mathcal{L}[X, \mathcal{A}, \mu]$*
2. *$\exists g \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ such that $|f| \leqslant_{a.e.} g$ and there exists a conegligible set $E \subseteq X$ such that $f_{|E}$ is $\mathcal{A}$-measurable.*
3. *$|f| \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ and there exists a conegligible set $E \subseteq X$ such that $f_{|E}$ is $\mathcal{A}$-measurable*

Proof.

1 ⇒ 2. As $f \in \mathcal{L}[X, \mathcal{A}]$ there exists $f_1 f_2 \in \mathcal{L}_+[X, \mathcal{A}]$ such that $f = f_1 - f_2$ [and thus $\operatorname{dom}(f) = \operatorname{dom}(f_1) \bigcap \operatorname{dom}(f2)$], so using 18.130 there exists conegligible sets E_1, E_2 such that $E_1 \subseteq \operatorname{dom}(f_1)$, $E_2 \subseteq \operatorname{dom}(f_2)$ and $f_{1|E_1}$, $f_{2|E_2}$ are $\mathcal{A}$-measurable. Take $E = E_1 \bigcap E_2$ then by 18.27 E is conegligible. As $f_{1|E_1}, f_{2|E_2}$ are $\mathcal{A}$-measurable we have using 18.103 that $f_{1|E} \underset{E \subseteq E_1 \text{ and } 2.28}{=} \left(f_{1|E_1}\right)_{|E}$, $f_{2|E} \underset{E \subseteq E_2 \text{ and } 2.28}{=} (f_{2|E_2})_{|E}$ are $\mathcal{A}$-measurable, and thus using 18.103 we have that

$$f_{|E} = f_{1|E} - f_{2|E} \text{ is } \mathcal{A}\text{-measurable} \tag{18.261}$$

Further as $f_1, f_2 \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ we have by 18.135 that $g = f_1 + f_2 \in \mathcal{L}_+[X, \mathcal{A}, \mu]$. Now $\forall x \in \operatorname{dom}(f)$ we have $|f(x)| = |f_1(x) - f_2(x)| \leqslant |f_1(x)| + |f_2(x)| = f_1(x) + f_2(x) = g(x)$ which as $\operatorname{dom}(f)$ is conegligible [see 18.137] proves that

$$|f| \leqslant_{a.e.} g \tag{18.262}$$

(2) is then proved by (18.261) and (18.262).

2 ⇒ 3. First by the hypothesis exists a conegligible set E such that $f_{|E}$ is $\mathcal{A}$-measurable. Using 18.103 (8) we have then that $|f_E|$ is $\mathcal{A}$-measurable. As $\forall x\in E$ we have $|f|_{|E}(x)=|f(x)|=|f_{|E}(x)|=|f_{|E}|(x)$ proving that $|f_E|=|f|_{|E}$ is $\mathcal{A}$-measurable. So we have

$$|\exists E \text{ congligible} \quad \text{such that } |f|_{|E} \text{ is } \mathcal{A}\text{-measurable} \tag{18.263}$$

As there exists also a $g\in\mathcal{L}_+[X,\mathcal{A},\mu]$, such that $|f|\leqslant_{a,e.} g$, it follows from 18.130 (2) together with $|f|(X)\subseteq[0,\infty[$ that

$$|f|\in\mathcal{L}_+[X,\mathcal{A},\mu] \tag{18.264}$$

3 ⇒ 1. Define $f^+=\frac{1}{2}\cdot(|f|+f)$ and $f^-=\frac{1}{2}\cdot(|f|-f)$ then using 18.142 we have

$$f^+(X), f^-(X)\subseteq[0,\infty[\tag{18.265}$$

Also $f^+(x)=\frac{1}{2}(|f(x)|+f(x))\leqslant\frac{1}{2}\cdot(|f(x)|+|f(x)|)=|f(x)|$ and $f^-(x)=\frac{1}{2}\cdot(|f(x)|-f(x))\underset{-f(x)\leqslant|f(x)|}{=}|f(x)|$ giving $\forall x\in\operatorname{dom}(f)=\operatorname{dom}(|f|)$ we have f^+, $f^-\leqslant|f|$ which as $\operatorname{dom}(f)$ is conegligible by 18.137 gives

$$f^+\leqslant_{a.e.}|f| \text{ and } f^-\leqslant_{a.e.}|f| \tag{18.266}$$

Further as $f_{|E}$ is $\mathcal{A}$-measurable we have by 18.103 (8) that $|f|_{|E}=|f_{|E}|$ is $\mathcal{A}$-measurable, applying 18.103 it follows that $f^+_{|E}=f^+=\frac{1}{2}\cdot(|f|_{|E}+f_{|E})$, $f^-_{|E}=\frac{1}{2}\cdot(|f|_{|E}+f_{|E})$ are $\mathcal{A}$-measurable, so

$$\exists E \text{ conegligible such that } f^+_{|E}, f^-_{|E} \text{ are } \mathcal{A}\text{-measurable} \tag{18.267}$$

As $|f|\in\mathcal{L}_+[X,\mathcal{A},\mu]$ by hypothesis we can use 18.130 (2) again to prove that $f^+,f^-\in\mathcal{L}_+[X.\mathcal{A},\mu]$ which as $f^+-f^-=\frac{1}{2}\cdot(|f|+f)-\frac{1}{2}\cdot(|f|-f)=f$ proves that

$$f\in\mathcal{L}[X,\mathcal{A},\mu]$$

□

Lemma 18.144. *Let $\langle X,\mathcal{A},\mu\rangle$ be a measure space, $f\in\mathcal{L}[X,\mathcal{A},\mu]$ such that $0\leqslant_{a.e.} f$ then there exists a conegligible set $E\in\mathcal{A}$ such that $E\subseteq\operatorname{dom}(f)$,$0\leqslant f_{|E}$, $f_{|E}$ is $\mathcal{A}$-measurable, $\forall a\in>0$ $\{a\leqslant f_{|E}\}\in\mathcal{A}$ and $\mu(\{a\leqslant f_{|E}\})<\infty$.*

Proof. First as $f\in\mathcal{L}[X,\mathcal{A},\mu]$ it follows from 18.143 that

$$|f|\in\mathcal{L}_+[X,\mathcal{A},\mu] \tag{18.268}$$

As $0\leqslant_{a.e.} f$ we have that $E_1=\{x\in\operatorname{dom}(f)|0\leqslant f(x)\}$ is conegligible. Further, as $\forall x\in E_1$ we have $f(x)=|f(x)|=|f|(x)$, it follows that $f_{|E_1}=_{a.e.}|f|$. Using 18.135 (4) it follows that $f_{|E_1}\in\mathcal{L}_+[X,\mathcal{A},\mu]$. Using 18.130 (1) there exists a conegligible set $E_2\in\mathcal{A}$ such that

$$E_2\subseteq\operatorname{dom}(f_{|E_1})\subseteq\operatorname{dom}(f)\wedge(f_{|E_1})_{|E_2} \text{ is } \mathcal{A}\text{-measurable and } \forall a>0\ \mu(\{a\leqslant(f_{|E_1})_{|E_2}\})<\infty \tag{18.269}$$

Using 18.27 there exists a conegligible set $E\in\mathcal{A}$ such that $E\subseteq E_1\bigcap E_2$. Now as $(f_{|E_1})_{E_2}$ is $\mathcal{A}$-measurable $((f_{|E_1})_{|E_2})_{|E}$ is $\mathcal{A}$-measurable [see 18.103 (7)]. Further we have $((f_{|E_1})_{|E_2})_{|E} \underset{2.28}{=} (f_{|E_1\bigcap E_2})_{|E} \underset{2.28}{=} f_{|E}$ so that

$$f_{|E} \text{ is } \mathcal{A}\text{-measurable}, 0\leqslant f_{|E} \text{ and } E\subseteq \operatorname{dom}(f) \text{ [as } E\subseteq E_2\subseteq \operatorname{dom}(f_{|E_1})\subseteq \operatorname{dom}(f)] \tag{18.270}$$

Using the above and the fact that $E\in\mathcal{A}$ allows us to apply 18.129 giving that

$$\forall a>0 \text{ we have } \{a\leqslant f_{|E}\}\in\mathcal{A} \tag{18.271}$$

As $\operatorname{dom}(f_{|E})=\operatorname{dom}(f)\bigcap E\subseteq\operatorname{dom}(f)\bigcap E_1\bigcap E_2=\operatorname{dom}(f_{|E_1\bigcap E_2}) \underset{2.28}{=} \operatorname{dom}((f_{|E_1})_{|E_2})$ we have that $\{a\leqslant f_{|E}\}=\{x\in\operatorname{dom}(f_{|E})|a\leqslant f_E(x)\}=\{x\in\operatorname{dom}(f)\bigcap E|a\leqslant f(x)\}\subseteq\{x\in\operatorname{dom}(f)\bigcap E_1\bigcap E_2|a\leqslant f(x)\}=\{x\in\operatorname{dom}((f_{|E_1})_{|E_2})|a\leqslant (f_{|E_1})_{|E_2}\}=\{a\leqslant (f_{|E_1})_{|E_1}\}$ it follows using 18.20 that

$$\mu(\{a\leqslant f_{|E}\})\leqslant\mu(\{a\leqslant (f_{|E_1})_{|E_2}\})\underset{(18.269)}{<}\infty \tag{18.272}$$ □

Corollary 18.145. *Let $\langle X,\mathcal{A},\mu\rangle$ be a measure space then*

1. *If $f\colon X\rightarrow\mathbb{R}$ is a non negative partial function then $f\in\mathcal{L}[X,\mathcal{A},\mu]\Leftrightarrow f\in\mathcal{L}_+[X,\mathcal{A},\mu]$*
2. *If $f\in\mathcal{L}[X,\mathcal{A},\mu]$ and $h\colon X\rightarrow\mathbb{R}$ a partial function such that $h=_{a.e.}f$ then $h\in\mathcal{L}[X,\mathcal{A},\mu]$ and $\int f d\mu=\int g d\mu$*
3. *If $f\in\mathcal{L}[X,\mathcal{A},\mu]$, $0\leqslant_{a.e.}f$ and $\int f d\mu\leqslant 0$ then $f=_{a.e.}0$*
4. *If $f\in[X,\mathcal{A},\mu]$, $f\leqslant_{a.e.}0$ and $0\leqslant\int f d\mu$ then $f=_{a.e.}0$*
5. *If $f,g\in\mathcal{L}[X,\mathcal{A},\mu]$ such that $f\leqslant_{a.e.}g$ and $\int g d\mu\leqslant\int f d\mu$ then $f=_{a.e.}g$*
6. *If $f\in\mathcal{L}[X,\mathcal{A},\mu]$ then $|f|\in\mathcal{L}[X,\mathcal{A},\mu]$ and $|\int f d\mu|\leqslant\int|f|d\mu$*

Proof.

1.

   ⇒. If $f\in\mathcal{L}[X,\mathcal{A},\mu]$ then by 18.143 $|f|\in\mathcal{L}_+[X,\mathcal{A},\mu]$. As f is non negative we have $f=|f|$ so that $f\in\mathcal{L}_+[X,\mathcal{A},\mu]$

   ⇐. This follows from 18.140.

2. As $f\in\mathcal{L}[X,\mathcal{A},\mu]$ we have by 18.143 that there exists a conegligible set $E\in\mathcal{A}$ such that

$$\exists g\in\mathcal{L}_+[X\mathcal{A},\mu] \text{ such that } |f|\leqslant_{a.e.}g \tag{18.273}$$

and

$$\exists E \text{ conegligible such that } f_{|E} \text{ is } \mathcal{A}\text{-measurable} \tag{18.274}$$

As $h=_{a.e.}f$ we have also that

$$F=\left\{x\in\operatorname{dom}(f)\bigcap\operatorname{dom}(h)|h(x)=f(x)\right\} \text{ is conegligible} \tag{18.275}$$

Now $\forall x \in E \bigcap F$ then $h_{|E \bigcap F}(x) = h(x) = f(x) = f_{|E \bigcap F}(x)$ which proves that $h_{|E \bigcap F} = f_{|E \bigcap F} \underset{2.28}{=} (f_{|E})_F$ which is $\mathcal{A}$-measurable because of (18.274) and 18.103 (7) so that

$$h_{|E \bigcap F} \text{ is } \mathcal{A}\text{-measurable and } E \bigcap F \text{ is conegligible} \tag{18.276}$$

From $|f| \leqslant_{a.e} g$ [see (18.273)] it follows that $G = \{x \in \operatorname{dom}(f) \bigcap \operatorname{dom}(g) \| f|(x) \leqslant g(x)\}$ is conegligible. Now if $x \in G \bigcap E \bigcap F$ a conegligible set [see 18.27] then $|h|(x) = |h(x)| = |f(x)| \leqslant g(x)$ which proves that

$$|h| \leqslant_{a.e} g \tag{18.277}$$

Applying 18.143 using (18.276) and (18.277) gives

$$h \in \mathcal{L}[X, \mathcal{A}, \mu]$$

Finally as $f =_{a.e} h$ implies $f \leqslant_{a.e.} h$ and $h \leqslant_{a.e.} f$ we have by 18.141 (4) that $\int f d\mu \leqslant \int h d\mu$ and $\int h d\mu \leqslant \int f d\mu$ giving

$$\int f d\mu = \int h d\mu$$

3. We prove this by contradiction, so assume that $f =_{a.e.} 0$ is false. As $f \in \mathcal{L}[X, \mathcal{A}, \mu]$ and $0 \leqslant_{a.e.} f$ it follows from 18.144 that there exists a conegligible set $E \in \mathcal{A}$ such that

$$E \subseteq \operatorname{dom}(f) \wedge 0 \leqslant f_{|E} \wedge f_{|E} \text{ is } \mathcal{A}\text{-measurable} \wedge \forall a > 0 \; \{a \leqslant f_{|E}\} \in \mathcal{A} \wedge \mu(\{a \leqslant f_{|E}\}) < \infty \tag{18.278}$$

Take now $F = \{0 < f_{|E}\}$ then as $E \in \mathcal{A}$, $E \subseteq \operatorname{dom}(f)$ we have by 18.129 that $F \in \mathcal{A}$ so that we $\mu(F)$ is well defined. Assume now that $\mu(F) = 0$ then F is by definition negligible so that $X \setminus F$ is conegligible, now if $x \in E \bigcap (X \setminus F)$ then $0 \leqslant f(x) \wedge 0 \nless f_{|E}(x) = f(x) \Rightarrow f(x) = 0$ so that $f =_{a.e.} 0$ contradicting the assumption that $f =_{a.e.} 0$ is not true. So we must have that

$$0 < \mu(F) \tag{18.279}$$

Given $n \in \mathbb{N}$ define then $F_n = \left\{\frac{1}{n} \leqslant f_{|E}\right\} = \left\{x \in \operatorname{dom}(f) \bigcap E | \frac{1}{n} \leqslant f_{|E}(x)\right\} \in \mathcal{A}$ [see (18.278)] then we have that

$$F = \bigcup_{n \in \mathbb{N}} F_n$$

Proof. If $x \in F$ then $0 < f_{|E}(x)$ then by 9.55 there exists a $n \in \mathbb{N}$ such that $0 < \frac{1}{n} < f_{|E}(x)$ so that $x \in F_n \subseteq \bigcup_{m \in \mathbb{N}} F_i$. Also if $x \in \bigcup_n F_n$ then there exists a $n \in \mathbb{N}$ such that $x \in F_n \Rightarrow 0 < \frac{1}{n} \leqslant f(x) \Rightarrow x \in F$. □

Next note that $0 < \mu(F) = \mu(\bigcup_{n \in \mathbb{N}} F_n) \leqslant \sum_{n=1}^{\infty} \mu(F_n)$ [see 18.20] so that $\exists n \in \mathbb{N}$ such that $0 < \mu(F_k)$. Define $g = \frac{1}{n} \cdot \mathcal{X}_{F_n}$ then using (18.278) we have that g is a simple function. Further $\forall x \in E$ we have either

$\boldsymbol{f_{|E}(x) < \frac{1}{n}}$**.** then $g(x) = \frac{1}{n} \cdot \mathcal{X}_{F_n}(x) = 0 \leqslant_{(18.278)} f_{|E}(x) = f(x)$

$\boldsymbol{\frac{1}{n} \leqslant f_{|E}(x)}$**.** then $g(x) = \frac{1}{n} \cdot \mathcal{X}_{F_n}(x) = \frac{1}{n} \leqslant f_{|E}(x) = f(x)$

which proves as E is conegligible that

$$g \leqslant_{a.e} f \tag{18.280}$$

Now as $g \in \mathcal{S}[X, \mathcal{A}, \mu] \subseteq_{18.133} \mathcal{L}_+[X, \mathcal{A}, \mu] \subseteq_{18.140} \mathcal{L}[X, \mathcal{A}, \mu]$ we have using the above and 18.141 (4) that

$$\int g d\mu \leqslant \int f d\mu. \tag{18.281}$$

Finally we have that $0 < \frac{1}{n} = \frac{1}{n} \cdot \mu(F_n) = \int g d\mu$ so that $0 < \int f d\mu$ which as we stated that $\int f d\mu \leqslant 0$ gives a contradiction. So the assumption that $f =_{a.e.} 0$ is false is wrong hence

$$f =_{a.e.} 0$$

4. Given $f \in \mathcal{L}[X, \mathcal{A}, \mu]$ we have using 18.141 (2) that $-f \in \mathcal{L}[X, \mathcal{A}, \mu]$. Further as $f \leqslant_{a.e.} 0$ we have that $\{x \in \operatorname{dom}(f) | f(x) \leqslant 0\}$ is μ-conegligible. Now we have

$$\begin{array}{rcl} x \in \{x \in \operatorname{dom}(f) | f(x) \leqslant 0\} & \Leftrightarrow & x \in \operatorname{dom}(f) \wedge f(x) \leqslant 0 \\ & \underset{\operatorname{dom}(f) = \operatorname{dom}(-f)}{\Leftrightarrow} & x \in \operatorname{dom}(-f) \wedge 0 \leqslant (-f)(x) \\ & \Leftrightarrow & x \in \{x \in \operatorname{dom}(-f) | 0 \leqslant (-f)(x)\} \end{array}$$

proving that $\{x \in \operatorname{dom}(-f) | 0 \leqslant (-f)(x)\}$ is μ-conegligible. Hence

$$(-f) \leqslant_{a.e.} 0 \tag{18.282}$$

Further $\int (-f) d\mu \underset{18.141\ (2)}{=} -\int f d\mu \leqslant 0$ [as by assumption $0 \leqslant \int f d\mu$]. We can now use (3) proving that $(-f) =_{a.e.} 0$ so that $\{x \in \operatorname{dom}(-f) | (-f)(x) = 0\}$ is μ-conegligible. As we have

$$\begin{array}{rcl} x \in \{x \in \operatorname{dom}(-f) | (-f)(x) = 0\} & \Leftrightarrow & x \in \operatorname{dom}(-f) \wedge -f(x) = 0 \\ & \underset{\operatorname{dom}(f) = \operatorname{dom}(-f)}{\Leftrightarrow} & x \in \operatorname{dom}(f) \wedge f(x) = 0 \\ & \Leftrightarrow & x \in \{x \in \operatorname{dom}(f) | f(x) = 0\} \end{array}$$

it follows that $\{x \in \operatorname{dom}(f) | f(x) = 0\}$ is μ-conegligible. So

$$f =_{a.e.} 0$$

5. From $f \leqslant_{a.e.} g$ it follows that $E = \{x \in \operatorname{dom}(f) \bigcap \operatorname{dom}(g) | f(x) \leqslant g(x)\}$ is conegligible. As $f(x) \leqslant g(x) \Leftrightarrow 0 \leqslant g(x) - f(x)$ we have that $0 \leqslant_{a.e.} g - f$. Further using 18.141 (1),(2),(3) we have that $f - g \in \mathcal{L}[X, \mathcal{A}, \mu]$ and as we stated that $\int g d\mu \leqslant \int f d\mu$ we have that $\int (g - f) d\mu = \int g d\mu - \int f d\mu \leqslant 0$. Applying then (3) proves that $g - f =_{a.e.} 0$ which as $g(x) - f(x) \Leftrightarrow (g - f)(x) = 0$ proves that

$$f =_{a.e.} g$$

6. As $f \in \mathcal{L}[X, \mathcal{A}, \mu]$ it follows from 18.143 that $|f| \in \mathcal{L}[X, \mathcal{A}, \mu]$. Define $f^+ = \frac{1}{2} \cdot (|f| + f)$ and $f^- = \frac{1}{2} \cdot (|f| - f)$ then we have using 18.141 that f^+, $f^- \in \mathcal{L}[X, \mathcal{A}, \mu]$ and using 18.142 that $0 \leqslant f^+, f^-$, which as $\operatorname{dom}(f^+) = \operatorname{dom}(f^-) = \operatorname{dom}(f)$ a conegligible set (see 18.137) proves that $0 \leqslant_{a.e.} f^+, f^-$ Now

$$\begin{aligned} \left|\int f d\mu\right| &\underset{18.142}{=} \left|\int (f^+ d\mu - f^-) d\mu\right| \\ &\underset{18.141}{=} \left|\int f^+ d\mu + \int f^- d\mu\right| \\ &\leqslant \left|\int f^+ d\mu\right| + \left|\int f^- d\mu\right| \\ &\leqslant_{0 \leqslant f^+, f^- \text{ and } 18.141(3)} \int f^+ d\mu + \int f^- d\mu \\ &\underset{18.141}{=} \int (f^+ + f^-) d\mu \\ &\underset{18.142}{=} \int |f| d\mu \end{aligned}$$

proving that

$$\left|\int f d\mu\right| \leqslant \int |f| d\mu$$

□

18.3 Limits and the measure integral

First we prove the Beppo Levi theorem that describes under which conditions the limit of a sequence of integrable functions is integrable and what its integral is. First we specify

Lemma 18.146. *Let $\langle X, \mathcal{A}, \mu \rangle$ be a measure space, $\{f_i\}_{i \in \mathbb{N}} \subseteq \mathcal{L}[X, \mathcal{A}, \mu]$ such that*

1. $f_1 =_{a.e.} 0$
2. $\forall n \in \mathbb{N}\ f_n \leqslant_{a.e.} f_{n+1}$
3. $\sup(\{\int f_n d\mu | n \in \mathbb{N}\}) < \infty$

then the partial function $\lim_{n \to \infty} f_n : X \to \mathbb{R}$ [see 18.88] is ***integrable*** *[in other words $\lim_{n \to \infty} f_n \in \mathcal{L}[X, \mathcal{A}, \mu]$] and $\int \left(\lim_{n \to \infty} f_n\right) d\mu = \lim_{n \to \infty} \int f_n d\mu = \sup(\{\int f_n d\mu | n \in \mathbb{N}\})$*

Proof. Let $n \in \mathbb{N}$ then from the fact that $f_n, f_{n+1} \in \mathcal{L}[X, \mathcal{A}, \mu]$ and $f_n \leqslant_{a.e.} f_{n+1}$ we have using 18.141 (4) that

$$\forall n \in \mathbb{N} \int f_n d\mu \leqslant \int f_{n+1} d\mu \tag{18.283}$$

so as $\sup(\{\int f_n d\mu | n \in \mathbb{N}\}) < \infty$ we have by 12.354

$$\lim_{n \to \infty} \int f_n d\mu = \sup\left(\left\{\int f_n d\mu | n \in \mathbb{N}\right\}\right) < \infty \tag{18.284}$$

From $f_1 =_{a.e.} 0$ we have by 18.30 that $0 \leqslant_{a.e.} f_1$ and from $f_n \leqslant_{a.e.} f_{n+1}$ we have by 18.30 that $f_1 \leqslant f_n$ and applying then 18.30 again gives

$$\forall n \in \mathbb{N} \text{ we have } 0 \leqslant_{a.e.} f_n \tag{18.285}$$

From the above and $f_n \in \mathcal{L}[X, \mathcal{A}, \mu]$ we have by 18.144 that

$\forall n \in \mathbb{N}$ there $\exists E_n \in \mathcal{A}$ such that E_n is conegligible, $E_n \subseteq \mathrm{dom}(f_n), 0 \leqslant f_{n|E_n}$, $f_{n|E_n}$ is $\mathcal{A}$-measurable and $\forall a > 0$ we have $\{a \leqslant f_{n|E_n}\} \in \mathcal{A}$ and $\mu(\{a \leqslant f_{n|E_n}\}) < \infty$ (18.286)

Further as $f_1 =_{a.e.} 0$ and $\forall n \in \mathbb{N}$ $f_n \leqslant_{a.e.} f_{n+1}$ we have by definition that

$F_0 = \{x \in \mathrm{dom}(f_1) | f_1(x) = 0\}$ and $\forall n \in \mathbb{N}\, F_n = \left\{x \in \mathrm{dom}(f_n) \bigcap \mathrm{dom}(f_{n+1}) | f_n(x) \leqslant f_{n+1}(x)\right\}$ are conegligible (18.287)

Using 18.27 we have that there exists a conegligible set E^* such that

$$E^* \in \mathcal{A} \text{ and } E^* \subseteq F_0 \bigcap \left(\bigcap_{n \in \mathbb{N}} F_n\right) \bigcap \left(\bigcap_{n \in \mathbb{N}} E_n\right) \tag{18.288}$$

Now as for $n \in \mathbb{N}$ $f_{n|E_n}$ is $\mathcal{A}$-measurable [see (18.286)] we have by 18.103 (7) that $(f_{n|E_n})_{|E^*}$ is $\mathcal{A}$-measurable and as $(f_{n|E_n})_{|E^*} \underset{E^* \subseteq E_n \text{ and } 2.28}{=} f_{n|E^*}$ we have

$$\forall n \in \mathbb{N} \text{ we have } f_{n|E^*} \text{ is } \mathcal{A}\text{-measurable} \tag{18.289}$$

Let $n \in \mathbb{N}$, $a > 0$ and define

$H_n(a) = \{a \leqslant f_{n|E^*}\} = \left\{x \in \mathrm{dom}(f_n) \bigcap E^* | a \leqslant f_n(x)\right\} \underset{E^* \subset \mathrm{dom}(f_n)}{=} \{x \in E^* | a \leqslant f(x)\}$ and $H(a) = \bigcup_{n \in \mathbb{N}} H_n(a)$ (18.290)

As $E^* \in \mathcal{A}$, $E^* \subseteq E_n \subseteq \mathrm{dom}(f_n)$ and $f_{n|E^*}$ is $\mathcal{A}$-measurable we have by 18.129 that $H_n(a) \in \mathcal{A}$. Further $H_n(a) = \{x \in \mathrm{dom}(f_n) \bigcap E^* | a \leqslant f_n(x)\} \subseteq \{x \in \mathrm{dom}(f_n) \bigcap E_n | a \leqslant f(x)\} = \{a \leqslant f_{n|E_n}\}$ so that $\mu(H_n(a)) \leqslant_{18.20} \mu(\{a \leqslant f_{n|H_n}\}) <_{(18.286)} \infty$ giving

$$\forall n \in \mathbb{N} \text{ we have } \mu(H_n(a)) \in \mathcal{A} \wedge \mu(H_n(a)) < \infty \tag{18.291}$$

Using the above we have that $a \cdot \mathcal{X}_{H_n(a)} \in \mathcal{S}[X, \mathcal{A}, \mu]$ and for $x \in E^* \subseteq E_n$ we have either

$f_n(x) < a$. then $a \cdot \mathcal{X}_{H_n(a)}(x) = 0 \leqslant_{(18.286)} f_n(x)$

$a \leqslant f_n(x)$. then $a \cdot \mathcal{X}_{H_n(a)}(x) = a \leqslant f_n(x)$

which proves as E^* is conegligible that $a \cdot \mathcal{X}_{H_n(a)} \leqslant_{a.e.} f_n$. Using this we have that $a \cdot \mu(H_n(a)) = \int^S a \cdot \mathcal{X}_{H_n(a)} d\mu \underset{18.133}{=} \int^+ a \cdot \mathcal{X}_{H_n(a)} d\mu \underset{18.140}{=} \int \alpha \cdot \mathcal{X}_{|H_n(a)} d\mu \leqslant_{18.141(4)} \int f_n d\mu \leqslant \sup(\{\int f_{\mathrm{n}\mathrm{d}} d\mu | n \in \mathbb{N}\})$ giving that

$$\forall a > 0 \text{ we have } \mu(H_n(a)) \leqslant \frac{I}{a} < \infty \text{ where } I = \sup\left(\left\{\int f_n d\mu | n \in \mathbb{N}\right\}\right) \tag{18.292}$$

Now $\forall n \in \mathbb{N}$ we have if $x \in H_n(a)$ then $x \in E^*$,$a \leqslant f_n(x)$ and as $f_n(x) \leqslant f_{n+1}(x)$ we have that $a \leqslant f_{n+1}(x)$ proving that $x \in H_{n+1}(a)$. So we have

$$\forall a > 0 \text{ we have } \forall n \in \mathbb{N} \text{ that } H_n(a) \subseteq H_{n+1}(a) \tag{18.293}$$

Using then 18.20 (5) gives

$$\forall a>0 \text{ we have } \mu(H(a)) = \lim_{n\rightarrow\infty} \mu(H_n(a)) = \sup(\{\mu(H_n(a))|n\in\mathbb{N}\}) \leqslant_{(18.292)} \frac{I}{a} < \infty \tag{18.294}$$

Next $\bigcap_{i\in\mathbb{N}} H(i)\in\mathcal{A}$ [by (18.291) and 18.8] and as $\bigcap_{i\in\mathbb{N}} H(i)\subseteq H(k)$ $\forall k\in\mathbb{N}$ we have that $\mu(\bigcap_{i\in\mathbb{N}} H(i))\leqslant\mu(H_k)\leqslant_{(18.294)}\frac{I}{k}$ so, as k is chosen arbitrary, $\mu(\bigcap_{i\in\mathbb{N}} H(i))\leqslant \inf\left(\{\frac{1}{k}|k\in\mathbb{N}\}\right) \underset{12.355}{=} 0$ proving that

$$\bigcap_{i\in\mathbb{N}} H(i) \text{ is negligible} \tag{18.295}$$

As E^* is conegligible, $E^*\in\mathcal{A}$ and $\forall n\in\mathbb{N}$ $H_n(i)\in\mathcal{A}$ [see (18.288),(18.291)] we have by 18.27 (6) and 18.8 that

$$E = E^*\setminus\left(\bigcap_{i\in\mathbb{N}} H(i)\right) \text{is conegligible and } E\in\mathcal{A} \tag{18.296}$$

If $x\in E$ there exists a $k\in\mathbb{N}$ such that $x\notin H(k)=\bigcup_{n\in\mathbb{N}} H_n(k)$ or $\forall n\in\mathbb{N}$ $x\notin H_n(k)\Rightarrow f_{n|E}(x)=f_n(x)<k$ so that by conditional completeness of $\mathbb{R}$ [see 9.43] $\sup(\{f_{n|E}(x)|n\in\mathbb{N}\})$ exists and is finite. Further as $0\leqslant f_n(x)$ [as $E\subseteq E^*\subseteq E^n$ and (18.286)] we have also that $0\leqslant\sup(\{f_{n|E}(x)|n\in\mathbb{N}\})$. As we have $f_{n|E}(x)=f_n(x)\leqslant f_{n+1}(x)=f_{n+1|E}(x)$ we have by 12.354 that $\sup(\{f_{n|E}(x)|n\in\mathbb{N}\})=\lim_{n\rightarrow\infty} f_{n|E}(x)=\lim_{n\rightarrow\infty} f_n(x)$. If $f=\lim_{n\rightarrow\infty} f_n$ [see 18.88] then we have $f_{|E}=\left(\lim_{n\rightarrow\infty} f_n\right)_{|E}=\lim_{n\rightarrow\infty} f_{n|E}$. To summarize we have

$$E\subseteq\mathrm{dom}(f),\ f_{|E}(X)\subseteq[0,\infty[\ f =_{a.e} f_{|E} \text{ where } f=\lim_{n\rightarrow\infty} f_n,\ f_{|E}=\left(\lim_{n\rightarrow\infty} f_n\right)_{|E} = \lim_{n\rightarrow\infty} f_{n|E} = \sup(\{f_{n|E}|n\in\mathbb{N}\}) \tag{18.297}$$

As $f_{n|E_n}$ is $\mathcal{A}$-measurable (see (18.286)) we have by 18.103 (7) that $(f_{n|E_n})_{|E}$ is $\mathcal{A}$-measurable and as $(f_{|E_n})_{|E} \underset{2.28}{=} f_{|E_n\bigcap E} \underset{E\subset E^*\subseteq E_n}{=} f_{n|E}$ it follows that $f_{n|E}$ is $\mathcal{A}$-measurable. Using then 18.104 we have that

$$f_{|E}=\lim_{n\rightarrow\infty} f_{n|E} \text{ is } \mathcal{A}\text{-measurable} \tag{18.298}$$

Next take $\varepsilon>0$ and $x\in\{\varepsilon\leqslant f_{|E}\}$ then $x\in E\subseteq E^\star$ and $\varepsilon\leqslant f_{|E}(x)=\sup(\{f_{n|E}(x)|n\in\mathbb{N}\})\Rightarrow\frac{\varepsilon}{2}<\sup(\{f_{n|E}(x)|n\in\mathbb{N}\})$, so there exist a $n\in\mathbb{N}$ such that $\frac{\varepsilon}{2}<f_{n|E}(x) \underset{E\subseteq E^*}{=} f_{n|E^*}(x)$ which proves that $\{\varepsilon\leqslant f_{|E}\}\subseteq\left\{\frac{\varepsilon}{2}\leqslant f_{n|E^*}\right\}=H_n\left(\frac{\varepsilon}{2}\right)\subseteq\bigcup_{i\in\mathbb{N}} H_i\left(\frac{\varepsilon}{2}\right)=H\left(\frac{\varepsilon}{2}\right)$. So

$$\{\varepsilon\leqslant f_{|E}\}\subseteq H\left(\frac{\varepsilon}{2}\right), \tag{18.299}$$

Further as $E\subseteq\mathrm{dom}(f)$, $E\in\mathcal{A}$, $f_{|E}$ is $\mathcal{A}$-measurable [see (18.296),(18.297) and (18.298)] we have by 18.129 that $\{\varepsilon\leqslant f_{|E}\}\in\mathcal{A}$. Next $\mu(\{\varepsilon\leqslant f_{|E}\})\leqslant_{(18.299),18.20}\mu\left(H\left(\frac{\varepsilon}{2}\right)\right)\leqslant_{(18.294)}\frac{2\cdot I}{\varepsilon}$. Hence we have

$$\forall\varepsilon>0\ \{\varepsilon\leqslant f_{|E}\}\in\mathcal{A}\wedge\mu(\{\varepsilon\leqslant f_{|E}\})\leqslant\frac{2\cdot I}{\varepsilon}<\infty \tag{18.300}$$

Let $g\in\mathcal{S}[X,\mathcal{A},\mu]$ be a simple function such that $g\leqslant_{a.e.}f_{|E}$. By 18.121 we have that

$$G=\{x\in X|g(x)\neq 0\}\in\mathcal{A}\wedge\mu(G)<\infty\wedge\exists M\in\mathbb{R}\vdash\forall x\in X g(x)\leqslant M \tag{18.301}$$

Let $\varepsilon>0$ then $\forall n\in\mathbb{N}$ set we have that $\mathrm{dom}(g-f_{n|E})=\mathrm{dom}(g)\bigcap\mathrm{dom}(f_{n|E})=X\bigcap E\bigcap\mathrm{om}(f_n)\underset{E\subseteq E_n\subseteq\mathrm{dom}(f_n)}{=}E$, further $g-f_{n|E}$ is $\mathcal{A}$-measurable [see 18.115,(18.298) and 18.103] so that by 18.129 $G_n(\varepsilon)=\{\varepsilon\leqslant g-f_{n|E}\}\in\mathcal{A}$. Now if $x\in G_{n+1}$ then $x\in E$ and $\varepsilon\leqslant(g-f_{n+1|E})(x)=g(x)-f_{n+1|E}(x)\leqslant g(x)-f_{n|E}(x)=(g-f_n)_{|E}(x)$ proving that $G_{n+1}(\varepsilon)\subseteq G_n(\varepsilon)$. So

$$\forall n\in\mathbb{N}\text{ we have }G_n(\varepsilon)\in\mathcal{A}\text{ and }G_{n+1}(\varepsilon)\subseteq G_n(\varepsilon)\text{ where }G_n(\varepsilon)=\{\varepsilon\leqslant g-f_{n|E}\} \tag{18.302}$$

Also if $x\in G_0(\varepsilon)$ then $x\in E$ and $0<\varepsilon\leqslant g(x)-f_{0|E}(x)\underset{E\subseteq E^*\subseteq F_0\text{ and }(18.287)}{=}g(x)$ so that $G_0(\varepsilon)\subseteq G$, hence we have $\mu(G_0(\varepsilon))\leqslant\mu(G)$, using (18.301) gives

$$\mu(G_0(\varepsilon))<\infty \tag{18.303}$$

(18.302) and (18.303) allows us to apply 18.20 (6) giving

$$\mu\left(\bigcap_{n\in\mathbb{N}}G_n(\varepsilon)\right)=\lim_{n\rightarrow\infty}\mu(G_n(\varepsilon))=\inf(\{\mu(G_n(\varepsilon))|n\in\mathbb{N}\}) \tag{18.304}$$

Now if $x\in\bigcap_{n\in\mathbb{N}}G_n(\varepsilon)$ then $x\in E$ and $\forall n\in\mathbb{N}$ we have that $\varepsilon\leqslant(g-f_{n|E})(x)=g(x)-f_n(x)$ so that $f_n(x)\leqslant g(x)-\varepsilon$, hence $f(x)=\sup(\{f_n(x)|n\in\mathbb{N}\})\leqslant g(x)-\varepsilon\Rightarrow f(x)<g(x)$ hence $x\notin\{x\in\mathrm{dom}(g)\bigcap\mathrm{dom}(f)|g(x)\leqslant f(x)\}$ proving that $\bigcap_{n\in\mathbb{N}}G_n(\varepsilon)\subseteq X\setminus\{x\in\mathrm{dom}(f)\bigcap\mathrm{dom}(g)|g(x)\leqslant f(x)\}$ which is negligible because $g\leqslant_{a.e.}f$. Hence we have that $\bigcap_{n\in\mathbb{N}}G_n(\varepsilon)$ is negligible so that $\mu(\bigcap_{n\in\mathbb{N}}G_n)=0$ which together with (18.304) gives that

$$\lim_{n\rightarrow\infty}\mu(G_n(\varepsilon))=\inf(\{\mu(G_n(\varepsilon))|n\in\mathbb{N}\})=0 \tag{18.305}$$

So by the above we have as $\varepsilon>0$ that $\inf(\{\mu(G_n(\varepsilon))\})<\varepsilon$ and there exists a $N_\varepsilon\in\mathbb{N}$ such that

$$0\leqslant\mu(G_{N_\varepsilon}(\varepsilon))<\varepsilon<\infty \tag{18.306}$$

Let now $x\in E\subseteq E^*\subseteq E_n$ then $f_{N_\varepsilon}(x)\geqslant 0$ [see (18.286)] and we have for $g(x)$ either

$g(x)\neq 0$. then $x\in G$, for $f_{|N_\varepsilon}(x)$ we have either

$\varepsilon\leqslant g(x)-f_{N_\varepsilon}(x)$. then $x\in G_{N_\varepsilon}(\varepsilon)$ so that

$$\begin{aligned}(f_n+\varepsilon\cdot\mathcal{X}_G+M\cdot\mathcal{X}_{G_{N_\varepsilon}(\varepsilon)})(x) &= f_{N_\varepsilon}(x)+\varepsilon\cdot\mathcal{X}_G(x)+M\cdot\mathcal{X}_{G_{N_\varepsilon}(\varepsilon)}(x)\\ &= f_{N_\varepsilon}(x)+\varepsilon+M\\ &\geqslant M\\ &\geqslant_{(18.301)} g(x)\end{aligned}$$

$\boldsymbol{g(x) - f_{N_\varepsilon}(x) < \varepsilon}$**.** then $x \notin G_{N_\varepsilon}(\varepsilon)$ so that

$$\begin{aligned}(f_n + \varepsilon \cdot \mathcal{X}_G + M \cdot \mathcal{X}_{G_{N_\varepsilon}(\varepsilon)})(x) &= f_{N_\varepsilon}(x) + \varepsilon \cdot \mathcal{X}_G(x) + M \cdot \mathcal{X}_{G_{N_\varepsilon}(\varepsilon)}(x) \\ &= f_{N_\varepsilon}(x) + \varepsilon \\ &\geqslant g(x)\end{aligned}$$

$\boldsymbol{g(x) = 0}$**.** then $x \notin G$ and for $f_{N_\varepsilon}(x)$ we have either

$\boldsymbol{\varepsilon \leqslant g(x) - f_{N_\varepsilon}(x)}$**.** then $x \in G_{N_\varepsilon}(\varepsilon)$ further as $g(x) = 0$ we have $0 \leqslant -f_{N_\varepsilon}(x) \Rightarrow f_{N_\varepsilon}(X) \leqslant 0$ proving that $f_{N_\varepsilon}(x) = 0$ so that

$$\begin{aligned}(f_n + \varepsilon \cdot \mathcal{X}_G + M \cdot \mathcal{X}_{G_{N_\varepsilon}(\varepsilon)})(x) &= f_N(x) + \varepsilon \cdot \mathcal{X}_G(x) + M \cdot \mathcal{X}_{G_n(\varepsilon)}(x) \\ &= 0 + M \\ &\geqslant_{(18.301)} g(x)\end{aligned}$$

$\boldsymbol{g(x) - f_{N_\varepsilon}(x) < \varepsilon}$**.** then $x \notin G_{N_\varepsilon}(\varepsilon)$

$$\begin{aligned}(f_n + \varepsilon \cdot \mathcal{X}_G + M \cdot \mathcal{X}_{G_{N_\varepsilon}(\varepsilon)})(x) &= f_N(x) + \varepsilon \cdot \mathcal{X}_G(x) + M \cdot \mathcal{X}_{G_{N_\varepsilon}(\varepsilon)}(x) \\ &= f_N(x) \\ &\geqslant 0 = g(x)\end{aligned}$$

The above proves as E is conegligible that

$$g \leqslant_{a.e.} f_{N_\varepsilon} + \varepsilon \cdot \mathcal{X}_G + M \cdot \mathcal{X}_{G_{N_\varepsilon}(\varepsilon)} \tag{18.307}$$

Using (18.303),(18.306) we have that $\varepsilon \cdot \mathcal{X}_G, M \cdot \mathcal{X}_{G_{N_\varepsilon}(\varepsilon)} \in \mathcal{S}[X, \mathcal{A}, \mu] \subseteq_{18.133} \mathcal{L}_+[X, \mathcal{A}, \mu] \subseteq_{18.140} \mathcal{L}[X, \mathcal{A}, \mu]$ so that

$$\begin{aligned}\int^S g d\mu &\underset{18.133, 18.140}{=} \int g d\mu \\ &\leqslant_{(18.307) \text{ and } 18.141} \int (f_{N_\varepsilon} + \varepsilon \cdot \mathcal{X}_G + M \cdot \mathcal{X}_{G_{N_\varepsilon}(\varepsilon)}) d\mu \\ &\underset{18.141}{=} \int f_{N_\varepsilon} d\mu + \int (\varepsilon \cdot \mathcal{X}_G) d\mu + \int (M \cdot \mathcal{X}_{G_{N_\varepsilon}(\varepsilon)}) d\mu \\ &\underset{18.133, 18.140}{=} \int f_{N_\varepsilon} d\mu + \int^S (\varepsilon \cdot \mathcal{X}_G) d\mu + \int^S (M \cdot \mathcal{X}_{G_{N_\varepsilon}(\varepsilon)}) d\mu \\ &= \int f_{N_\varepsilon} d\mu + \varepsilon \cdot \mu(G) + M \cdot \mu(G_{N_\varepsilon}(\varepsilon)) \\ &\leqslant \sup\left(\left\{\int f_n d\mu | n \in \mathbb{N}\right\}\right) + \varepsilon \cdot \mu(G) + M \cdot \mu(G_{N_\varepsilon}(\varepsilon)) \\ &\leqslant_{(18.306)} \sup\left(\left\{\int f_n d\mu | n \in \mathbb{N}\right\}\right) + \varepsilon \cdot \mu(G) + M \cdot \varepsilon \\ &= \sup\left(\left\{\int f_n d\mu | n \in \mathbb{N}\right\}\right) + \varepsilon \cdot (\mu(G) + M)\end{aligned}$$

As $g \in \mathcal{S}[X, \mathcal{A}, \mu]$ and $\varepsilon > 0$ are chosen arbitrary we have [see 9.56]

$$\forall g \in \mathcal{S}[X, \mathcal{A}, \mu] \text{ with } g \leqslant_{a.e.} f_{|E} \text{ we have} \int^S g d\mu = \int g d\mu \leqslant \sup\left(\left\{\int f_n d\mu | n \in \mathbb{N}\right\}\right) < \infty \tag{18.308}$$

So $\sup\left(\left\{\int^S g\, d\mu | g \in \mathcal{S}[X, \mathcal{A}, \mu] \wedge g \leqslant_{a.e.} f_{|E}\right\}\right) < \infty$, further we have $E \subseteq_{(18.297)} \operatorname{dom}(f) \Rightarrow E = \operatorname{dom}(f_{|E})$, $f_{|E}$ is $\mathcal{A}$-measurable [see (18.298)] and $\forall \varepsilon > 0\, \mu(\{\varepsilon \leqslant f_{|E}\}) < \infty$ [see (18.300)]. These are the conditions for lemma 18.130 and applying this gives

$$f_{|E} \in \mathcal{L}_+[X, \mathcal{A}.\mu] \subseteq_{18.140} \mathcal{L}[X, \mathcal{A}, \mu] \tag{18.309}$$

Further

$$\begin{aligned}
\int f_{|E} d\mu &\underset{18.140}{=} \int^+ f_{|E} d\mu \\
&\underset{\text{def}}{=} \sup\left(\left(\left\{\int^S g d\mu | g \in \mathcal{S}[X, \mathcal{A}, \mu] \wedge g \leqslant_{a.e.} f_{|E}\right\}\right)\right) \\
&\leqslant_{(18.308)} \sup\left(\left\{\int f_n d\mu | n \in \mathbb{N}\right\}\right)
\end{aligned}$$

Now as E is conegligible $f =_{a.e.} f_{|E}$ so using 18.145 (2) on (18.309) and the above we have

$$\lim_{n \to \infty} f_n = f \in \mathcal{L}[X, \mathcal{A}, \mu] \text{ and } \int f d\mu \leqslant \sup\left(\left\{\int f_n d\mu | n \in \mathbb{N}\right\}\right) \tag{18.310}$$

Now as $\forall n \in \mathbb{N}\ f_n \leqslant_{a.e.} f$ [see (18.297) + E is conegligible] we have by 18.141 (4)] that $\int f_n d\mu \leqslant \int f d\mu$, so $\sup(\{\int f_n d\mu | n \in \mathbb{N}\}) \leqslant \int f d\mu$ which combined with (18.310) gives finally

$$\lim_{n \to \infty} f_n \in \mathcal{L}[X, \mathcal{A}, \mu] \text{ and } \int \left(\lim_{n \to \infty} f_n\right) d\mu = \sup\left(\left\{\int f_n d\mu | n \in \mathbb{N}\right\}\right) \underset{(18.284)}{=} \lim_{n \to \infty} \int f_n d\mu$$

□

We use now the above lemma to prove the general case of Beppo Levi

Theorem 18.147. (Beppo Levi) *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space, $\{f_i\}_{i \in \mathbb{N}} \subseteq \mathcal{L}[X, \mathcal{A}, \mu]$ such that*

1. $\forall n \in \mathbb{N}\ f_n \leqslant_{a.e.} f_{n+1}$
2. $\sup(\{\int f_n d\mu | n \in \mathbb{N}\}) < \infty$

*then the partial function $\lim_{n \to \infty} f_n: X \to \mathbb{R}$ [see 18.88] is **integrable** [in other words $\lim_{n \to \infty} f_n \in \mathcal{L}[X, \mathcal{A}, \mu]$] and $\int \left(\lim_{n \to \infty} f_n\right) d\mu = \lim_{n \to \infty} \int f_n d\mu$*

Proof. Define $\{h_n\}_{n \in \mathbb{N}}$ by $h_n = f_n - f_1$ then by 18.141

$$\{h_n\}_{n \in \mathbb{N}} \subseteq \mathcal{L}[X, \mathcal{A}, \mu]. \tag{18.311}$$

Then as $\mathrm{dom}(f_1)$ is conegligible $[f_0 \in \mathcal{L}[X, \mathcal{A}, \mu] + 18.137]$ and $\forall x \in \mathrm{dom}(f)\ h_1(x) = f_1(x) - f_1(x) = 0$ we have

$$0 =_{a.e.} h_0 \tag{18.312}$$

Let $n \in \mathbb{N}$ then $\mathrm{E} = \{x \in \mathrm{dom}(f_n) \bigcap \mathrm{dom}(f_{n+1}) | f_n(x) \leqslant f_{n+1}(x)\}$ is a conegligible set and as $f_1 \in \mathcal{L}[X, \mathcal{A}, \mu]$ $\mathrm{dom}(f_1)$ is conegligible [see 18.137] we have that $E \bigcap \mathrm{dom}(f_1)$ is conegligible. If $x \in E \bigcap \mathrm{dom}(f_1)$ we have $h_n(x) = f_n(x) - f_1(x) \leqslant f_{n+1}(x) - f_1(x) = h_{n+1}(x)$ proving

$$\forall n \in \mathbb{N} \text{ we have } h_n \leqslant h_{n+1} \tag{18.313}$$

Next as $\forall n \in \mathbb{N}$we have $\int f_n d\mu \leqslant \sup(\{\int f_n d\mu | n \in \mathbb{N}\})$ we have $\int h_n d\mu = \int f_n d\mu - \int f_1 d\mu \leqslant \sup(\{\int f_n d\mu | n \in \mathbb{N}\}) - \int f_1 d\mu$ so that

$$\sup\left(\left\{\int h_n d\mu | n \in \mathbb{N}\right\}\right) \leqslant \sup\left(\left\{\int f_n d\mu | n \in \mathbb{N}\right\}\right) - \int f_1 d\mu < \infty \tag{18.314}$$

(18.311),(18.312),(18.313) and (18.314) allows us to use the previous theorem 18.146 giving that

$$\lim_{n\rightarrow\infty} h_n \in \mathcal{L}[X, \mathcal{A}, \mu] \text{ and } \int \left(\lim_{n\rightarrow\infty} h_n\right) d\mu = \lim_{n\rightarrow\infty} \int h_n d\mu \tag{18.315}$$

Now as $\mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right) \in \mathcal{L}[X, \mathcal{A}, \mu]$ we have by 18.137 that $\mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right)$ is conegligible, as also $\mathrm{dom}(f_1)$ is conegligible we have that $\mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right) \bigcap \mathrm{dom}(f_1)$ is conegligible. Let $x \in \mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right) \bigcap \mathrm{dom}(f_1)$ then $\left(\lim_{n\rightarrow\infty} h_n + f_1\right)(x) = \left(\lim_{n\rightarrow\infty} h_n(x)\right) + f_1(x) \underset{\text{Algorithm } 12.340}{=} \lim_{n\rightarrow\infty} (h_n(x) + f_1(x)) = \lim_{n\rightarrow\infty} f_n(x)$ proving that

$$\lim_{n\rightarrow\infty} h_n + f_1 =_{a.e} \lim_{n\rightarrow\infty} f_n \tag{18.316}$$

Using 18.141 we have that $\lim_{n\rightarrow\infty} h_n + f_1 \in \mathcal{L}[X, \mathcal{A}, \mu]$ so that by 18.145 (2) we have that

$$\lim_{n\rightarrow\infty} f_n \in \mathcal{L}[X, \mathcal{A}, \mu]$$

Further

$$\begin{aligned}
\int \lim_{n\rightarrow\infty} f_n &\underset{18.145}{=} \int \left(\lim_{n\rightarrow\infty} h_n\right) d\mu + \int f_1 d\mu \\
&\underset{(18.315)}{=} \lim_{n\rightarrow\infty} \int h_n d\mu + \int f_1 d\mu \\
&= \lim_{n\rightarrow\infty} \int (f_n - f_1) d\mu + \int f_1 d\mu \\
&\underset{18.141}{=} \lim_{n\rightarrow\infty} \left(\int f_n d\mu - \int f_1 d\mu\right) + \int f_1 d\mu \\
&\underset{\text{Algorithm } 12.340}{=} \left(\lim_{n\rightarrow\infty} \int f_n d\mu\right) - \int f_1 d\mu + \int f_1 d\mu \\
&= \lim_{n\rightarrow\infty} \int f_n d\mu
\end{aligned}$$

which proves the theorem. □

Lemma 18.148. (Fatou's Lemma) *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space, $\{f_n\}_{n\in\mathbb{N}}\subseteq \mathcal{L}[X, \mathcal{A}, \mu]$ such that*

1. $\forall n\in\mathbb{N}\ 0\leqslant_{a.e.} f_n$
2. $\liminf_{n\to\infty}\int f_n<\infty$

then $\liminf_{n\to\infty} f_n\in\mathcal{L}[X, \mathcal{A}, \mu]$ *[see 18.88] and* $\int\left(\liminf_{n\to\infty} f_n\right)d\mu\leqslant\liminf_{n\to\infty}\int f_n d\mu$

Proof. Let $n\in\mathbb{N}$ then as $f_n\in\mathcal{L}[X, \mathcal{A}, \mu]$ and $0\leqslant_{a.e.} f_n$ we can apply 18.144 giving

$$\exists E_n\in\mathcal{A} \text{ with } E_n \text{ conegligble, } E_n\subseteq \operatorname{dom}(f_n),\ 0\leqslant f_{n|E_n} \text{ and } f_{n|E_n} \text{ is } \mathcal{A}\text{-measurable} \tag{18.317}$$

Define now $\{g_n\}_{n\in\mathbb{N}}$ by $g_n=\inf_{i\in\mathbb{N}} f_{n+i-1|E}$ then we have using the above together with 18.104 that

$$\forall n\in\mathbb{N}\, g_n \text{ is } \mathcal{A}\text{-measurable} \tag{18.318}$$

Further as $\forall x\in\bigcap_{i\in\mathbb{N}} E_{n+i-1}$ we have that $\{f_{n+i-1|E_{n+i-1}}(x)|i\in\mathbb{N}\}$ is bounded below by 0 we have by 9.43 that $\left(\inf_{i\in\mathbb{N}} f_{n+i-1|E_{n+i-1}}\right)(x)=\inf(\{f_{n+i-1|E_{n+i-1}}(x)|i\in\mathbb{N}\})$ exists and $0\leqslant\left(\inf_{i\in\mathbb{N}} f_{n+i-1|E_{|n+i-1}}\right)(x)$ so we have taking in account 18.88 (10) that

$$\forall n\in\mathbb{N} \text{ we have } \operatorname{dom}(g_n)\subseteq\bigcap_{i\in\mathbb{N}} E_{n+i-1} \text{ [conegligble by (18.317),18.27] } \underset{18.25}{\Rightarrow} \operatorname{dom}(f_n) \text{ is conegligible and } 0\leqslant g_n \tag{18.319}$$

Further $\forall x\in\operatorname{dom}(g_n)$ we have $|g_n|(x)=g_n(x)=\inf(\{f_{i+n-1|E_{n+i-1}}(x)|i\in\mathbb{N}\})\leqslant f_{n|E_n}(x)\underset{(18.317)}{=}|f_n(x)|$ which as $\operatorname{dom}(g_n)$ is conegligible means that

$$|g_n|\leqslant_{a.e.}|f_n|$$

As $f_n\in\mathcal{L}[X, \mathcal{A}, \mu]$ we have by 18.143 (3) that $|f_n|\in\mathcal{L}_+[X, \mathcal{A}, \mu]$, applying then 18.143 (2) prove then that

$$\forall n\in\mathbb{N} \text{ we have that } g_n\in\mathcal{L}[X, \mathcal{A}, \mu] \text{ or } \{g_n\}_{n\in\mathbb{N}}\subseteq\mathcal{L}[X, \mathcal{A}, \mu] \tag{18.320}$$

Take $n\in\mathbb{N}$ and $m\geqslant n$ let $i=m-n+1\in\mathbb{N}$ so that $m=n+i-1$ then $\forall x\in\operatorname{dom}(g_n)$ we have $g_n(x)=\inf(\{f_{i+n-1|E_{n+i-1}}(x)|i\in\mathbb{N}\})\leqslant f_{m|E}(x)=f_m(x)$ which as $\operatorname{dom}(g_n)$ is conegligible means that $g_n\leqslant_{a.e.} f_m$ so that by 18.141 $\int g_n d\mu\leqslant\int f_m d\mu$, hence $\int g_n d\mu\leqslant\inf(\int f_m d\mu|m\in\{n,\ldots,\infty\})$. So $\sup(\int g_n d\mu|n\in\mathbb{N})\leqslant\sup(\{\inf(\int f_m d\mu|m\in\{n,\ldots,\infty\})|n\in\mathbb{N}\})=\liminf_{n\to\infty}\int f_n d\mu<\infty$ proving that

$$\sup\left(\left\{\int g_n d\mu|n\in\mathbb{N}\right\}\right)\leqslant\liminf_{n\to\infty}\int f_n d\mu<\infty \tag{18.321}$$

Further let $n \in \mathbb{N}$ then for $x \in \mathrm{dom}(g_n) \bigcap \mathrm{dom}(g_{n+1})$ [a conegligible set because (18.319) and 18.27] then $\{f_{i+(n+1)-1|E_{n+i-1}}(x)|i \in \mathbb{N}\} \subseteq \{f_{i+n-1|E_{n+i-1}}(x)|i \in \mathbb{N}\}$ so that $g_n(x) = \inf(\{f_{i+n-1|E_{n+i-1}}(x)|i \in \mathbb{N}\}) \leqslant \inf(\{f_{i+(n+1)-1|E_{n+i-1}}(x)|i \in \mathbb{N}\}) = g_{n+1}(x)$ proving that

$$\forall n \in \mathbb{N} \text{ we have } g_n \leqslant_{a.e.} g_{n+1} \tag{18.322}$$

(18.320),(18.322) and (18.321) allows use to use Beppo Levi (see 18.147) giving

$$\lim_{n\rightarrow\infty} g_n \in \mathcal{L}[X, \mathcal{A}, \mu] \text{ and } \int \left(\lim_{n\rightarrow\infty} g_n\right) d\mu = \lim_{n\rightarrow\infty} \int g_n d\mu \tag{18.323}$$

Now if $x \in \mathrm{dom}\left(\lim_{n\rightarrow\infty} g_n\right)$ [a conegligible set because $\lim_{n\rightarrow\infty} g_n \in \mathcal{L}[X, \mathcal{A}, \mu]$ and 18.137] we have that

$$\begin{aligned}
\left(\lim_{n\rightarrow\infty} g_n\right)(x) &= \lim_{n\rightarrow\infty} g_n(x) \\
&\underset{(18.322) \text{ and } 12.354}{=} \sup(\{g_n(x)|n \in \mathbb{N}\}) \\
&= \sup(\{\inf(\{f_{n+i-1|E_{n+i-1}}(x)|i \in \mathbb{N}\})|n \in \mathbb{N}\}) \\
&= \sup(\{\inf(\{f_i(x)_{|E_i}|i \in \{n, \ldots, \infty\}\})|n \in \mathbb{N}\}) \\
&= \sup(\{\inf(\{f_i(x)|i \in \{n, \ldots, \infty\}\})|n \in \mathbb{N}\}) \\
&= \liminf_{n\rightarrow\infty} f_n(x)
\end{aligned}$$

proving that $\lim_{n\rightarrow\infty} g_n =_{a.e.} \liminf_{n\rightarrow\infty} f_n$, using then 18.145 (2) on (18.323) produces

$$\liminf_{n\rightarrow\infty} f_n \in \mathcal{L}[X, \mathcal{A}, \mu] \text{ and } \int \left(\liminf_{n\rightarrow\infty} f_n\right) d\mu = \lim_{n\rightarrow\infty} \int g_n d\mu \tag{18.324}$$

Now applying 18.141 (4) on (18.322) gives that $\forall n \in \mathbb{N}$ $\int g_n d\mu \leqslant \int g_{n+1} d\mu$ so that by 12.354 we have that $\lim_{n\rightarrow\infty} \int g_n d\mu = \sup(\{\int g_n d\mu | n \in \mathbb{N}\})$. Applying then (18.321) on (18.324) gives

$$\int \left(\liminf_{n\rightarrow\infty} f_n\right) d\mu \leqslant \liminf_{n\rightarrow\infty} \int f_n d\mu \tag{18.325}$$

The lemma is then proved by (18.324) and (18.325). □

Theorem 18.149. (Lebesgue dominant convergence) *Let $\langle X, \mathcal{A}, \mu \rangle$ be a measure space, $\{f_n\}_{n \in \mathbb{N}} \subseteq \mathcal{L}[X, \mathcal{A}, \mu]$ such that $\mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right)$ is conegligible [see 18.88] and there exists a $g \in \mathcal{L}[X, \mathcal{A}, \mu]$ with $\forall n \in \mathbb{N}$ $|f_n| \leqslant_{a.e.} g$ then $\lim_{n\rightarrow\infty} f_n \in \mathcal{L}[X, \mathcal{A}, \mu]$ and $\int \left(\lim_{n\rightarrow\infty} f_n\right) d\mu = \lim_{n\rightarrow\infty} \int f_n d\mu$*

Proof. Let $n \in \mathbb{N}$ and define $h_n^+ = f_n + g$ and $h_n^- = (-f_n) + g$ then by 18.141 we have that

$$\{h_n^+\}_{n \in \mathbb{N}} \subseteq \mathcal{L}[X, \mathcal{A}, \mu] \wedge \{h_n^-\}_{n \in \mathbb{N}} \subseteq \mathcal{L}[X, \mathcal{A}, \mu] \tag{18.326}$$

As $\forall n \in \mathbb{N}$ $|f_n| \leqslant_{a.e.} g$ we have that $E_n = \{x \in \mathrm{dom}(f_n) \bigcap \mathrm{dom}(g) | |f_n(x)| \leqslant g(x)\}$ is conegligible and for $x \in E_n$ we have that

1. $h_n^+(x) = f_n(x) + g(x) \leqslant |f_n(x)| + g(x) \leqslant g(x) + g(x) = 2 \cdot g(x)$
2. $h_n^+(x) = f_n(x) + g(x) \geqslant f_n(x) + |f_n(x)| \geqslant -|f(x)| + |f(x)| = 0$

3. $h_n^-(x) = -f_n(x) + g(x) \leqslant |f_n(x)| + g(x) \leqslant g(x) + g(x) = 2 \cdot g(x)$
4. $h_n^-(x) = -f_n(x) + g(x) \geqslant -f_n(x) + |f_n(x)| \geqslant -|f(x)| + |f(x)| = 0$

proving that

$$\forall n \in \mathbb{N} \text{ we have } 0 \leqslant_{a.e.} h_n^+ \leqslant_{a.e.} 2 \cdot g \wedge 0 \leqslant_{a.e} h_n^- \leqslant_{a.e.} 2 \cdot g \tag{18.327}$$

Next using 18.141 (4) we have $\forall n \in \mathbb{N}$ that $0 \leqslant \int h_n^+ d\mu \leqslant 2 \cdot \int g d\mu$, $0 \leqslant \int h_n^- d\mu \leqslant 2 \cdot \int g d\mu$ so that for $m \in \{n, \ldots, \infty\}$ we have by the conditional completeness of $\mathbb{R}$ [see 9.43] that the the following infinum's exists and that $\inf(\{\int h_i^+ | i \in \{m, \ldots, \infty\}\}) \leqslant 2 \cdot \int g d\mu$, $\inf(\{\int h_i^- | i \in \{m, \ldots, \infty\}\}) \leqslant 2 \cdot \int g d\mu$. Using the conditional completeness of $\mathbb{R}$ again the following supremum's exists $\sup(\{\inf(\{\int h_i^+ | i \in \{m, \ldots, \infty\}\}) | m \in \mathbb{N}\}) \leqslant 2 \cdot \int g d\mu < \infty$, $\sup(\{\inf(\{\int h_i^- | i \in \{m, \ldots, \infty\}\}) | m \in \mathbb{N}\}) \leqslant 2 \cdot \int g d\mu < \infty$ proving that

$$\liminf_{n \rightarrow \infty} \int h_n^+ d\mu < \infty \wedge \liminf_{n \rightarrow \infty} \int h_n^- d\mu < \infty \tag{18.328}$$

Using (18.326) and (18.328) we can apply Fatou's lemma [see 18.148] proving that

$$h^+ = \liminf_{n \rightarrow \infty} h_n^+, h^- = \liminf_{n \rightarrow \infty} h_n^- \in \mathcal{L}[X, \mathcal{A}, \mu] \tag{18.329}$$

and

$$\int h^+ d\mu = \leqslant \liminf_{n \rightarrow \infty} \int h_n^+ d\mu, \int h^- d\mu \leqslant \liminf_{n \rightarrow \infty} \int h_n^- d\mu \tag{18.330}$$

Take now $f = \lim_{n \rightarrow \infty} f_n$ then $\forall x \in \operatorname{dom}(f)$ we have $f(x) = \lim_{n \rightarrow \infty} f_x(x)$ then we have by 12.360 that

$$\forall x \in \operatorname{dom}(f) \text{ we have } f(x) = \lim_{n \rightarrow \infty} f_n(x) = \liminf_{n \rightarrow \infty} f_n(x) = \limsup_{n \rightarrow \infty} f_n(x) \tag{18.331}$$

Now as $g, h^+, h^- \in \mathcal{L}[X, \mathcal{A}, \mu]$ we have by 18.137 that $\operatorname{dom}(g), \operatorname{dom}(h^+), \operatorname{dom}(h^-)$ are conegligible and we stated that $\operatorname{dom}(f)$ is conegligible so that [using 18.27]

$$\operatorname{dom}(f) \bigcap \operatorname{dom}(g) \bigcap \operatorname{dom}(h^+), \operatorname{dom}(f) \bigcap \operatorname{dom}(g) \bigcap \operatorname{dom}(h^-) \text{ are conegligible} \tag{18.332}$$

Now $\forall x \in \operatorname{dom}(f) \bigcap \operatorname{dom}(g) \bigcap \operatorname{dom}(h^+)$ $\liminf_{n \rightarrow \infty} h_n^+(x)$ exists so using 12.362 $\liminf_{n \rightarrow \infty} (h_n^+(x) - g(x))$ exists and $\liminf_{n \rightarrow \infty} (h_n^+(x) - g(x)) = \liminf_{n \rightarrow \infty} h_n^+(x) - g(x)$. As $\liminf_{n \rightarrow \infty} (h_n^+(x) - g(x)) = \liminf_{n \rightarrow \infty} (h_n^+(x) - g(x)) = \liminf_{n \rightarrow \infty} f_n(x) \underset{(18.331)}{=} f(x)$ we have that

$$f =_{a.e.} h^+ - g \tag{18.333}$$

Likewise $\forall x \in \operatorname{dom}(f) \bigcap \operatorname{dom}(g) \bigcap \operatorname{dom}(h^-)$ $\liminf_{n \rightarrow \infty} h_n^-(x)$ exists so using 12.362 $\liminf_{n \rightarrow \infty} (h_n^-(x) - g(x))$ exists and $\liminf_{n \rightarrow \infty} (h_n^-(x) - g(x)) = \liminf_{n \rightarrow \infty} h_n^-(x) - g(x)$. As $\liminf_{n \rightarrow \infty} (h_n^-(x) - g(x)) = \liminf_{n \rightarrow \infty} (h_n^-(x) - g(x)) = \liminf_{n \rightarrow \infty} (-f_n(x)) \underset{17.72}{=} -\limsup_{n \rightarrow \infty} f_n(x) \underset{(18.331)}{=} -f(x)$ we have that

$$-f =_{a.e.} h^- - g \tag{18.334}$$

Using (18.329) with 18.141 we have that $h^+ - g, h^- - g \in \mathcal{L}[X, \mathcal{A}, \mu]$, combining then (18.333),(18.334) with 18.145 (2) proves that

$$f, -f \in \mathcal{L}[X, \mathcal{A}, \mu] \text{ and } \int f d\mu = \int h^+ d\mu - \int g d\mu, \int (-f) d\mu = \int h^- d\mu - \int g d\mu \tag{18.335}$$

Hence

$$\begin{aligned}
\int f d\mu &= \int h^+ d\mu - \int g d\mu \\
&\leqslant_{(18.330)} \liminf_{n\to\infty} \int h_n^+ d\mu - \int g d\mu \\
&\underset{12.362}{=} \liminf_{n\to\infty} \int (h_n^+ - g) d\mu \\
&= \liminf_{n\to\infty} \int f_n d\mu \qquad (18.336) \\
-\int f d\mu &= \int (-f) d\mu \\
&= \left(\int h^- d\mu - \int g d\mu \right) \\
&\leqslant_{(18.330)} \left(\liminf_{n\to\infty} \int h_n^- d\mu - \int g d\mu \right) \\
&\underset{12.362}{=} \liminf_{n\to\infty} \int (h_n^- - g) d\mu \\
&= \liminf_{n\to\infty} \int (-f_n) d\mu \\
&= \liminf_{n\to\infty} \left(-\int f_n d\mu \right) \\
&\underset{17.72}{=} -\limsup_{n\to\infty} \int f_n d\mu \qquad (18.337)
\end{aligned}$$

So we have that $\int f d\mu \leqslant_{(18.336)} \liminf_{n\to\infty} \int f_n d\mu \leqslant_{12.359} \limsup_{n\to\infty} \int f_n d\mu \leqslant_{(18.337)} \int f d\mu$ proving that $\int f d\mu = \liminf_{n\to\infty} \int f_n d\mu = \limsup_{n\to\infty} \int f_n d\mu$. Applying then 12.360 gives

$$\int f d\mu = \lim_{n\to\infty} \int f_n d\mu \tag{18.338}$$

The first part of (18.335) and the above proves then the theorem. □

Up to now we have assumed that $f(X) \subseteq \mathbb{R}$ [or f is finite] we can easely extend our definition of the measure integral to the inifinite case as follows.

Definition 18.150. *Let $\langle X, \mathcal{A}, \mu \rangle$ be a measure space and $f: X \to \bar{\mathbb{R}}$ then $f \in \bar{\mathcal{L}}[X, \mathcal{A}, \mu]$ if and only if $\bar{f} \in \mathcal{L}[X, \mathcal{A}, \mu]$ where $\bar{f}: X \to \mathbb{R}$ is the partial defined by $\bar{f} = f_{|\{x \in \mathrm{dom}(f) | f(x) \in \mathrm{Rr}\}}$ do $\mathrm{dom}(\bar{f}) = \{x \in \mathrm{dom}(f) | f(x) \in \mathbb{R}\}$. If $\bar{f} \in \bar{\mathcal{L}}[X, \mathcal{A}, \mu]$ then $\bar{\int} f d\mu = \int \bar{f} d\mu$*

Remark 18.151. Let $\langle X, \mathcal{A}, \mu\rangle$ and $f \in \bar{\mathcal{L}}[X, \mathcal{A}, \mu]$ then $\operatorname{dom}(f)$ is μ-conegligible

Proof. By definition we have $\bar{f} \in \mathcal{L}[X, \mathcal{A}, \mu]$ so that by 18.137 $\operatorname{dom}(\bar{f})$ is μ-conegligible which as $\operatorname{dom}(\bar{f}) \subseteq \operatorname{dom}(f)$ proves by 18.27 that $\operatorname{dom}(f)$ is μ-conegligible. □

18.4 Integration over sub-spaces

Proposition 18.152. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space and $E \in \mathcal{A}$ then for the measure space $\langle E, \mathcal{A}_E, \mu_E\rangle$ [see 18.36] we have*

1. $\mathcal{A}_E = \{A \in \mathcal{A} | A \subseteq E\} \subseteq \mathcal{A}$
2. $\mu_E: \mathcal{A}_E \rightarrow [0, \infty]$ *is defined by* $\mu_E = \mu_{|\mathcal{A}_E}$

Proof.

1. First note that by 18.36 $\mathcal{A}_E = \{A \bigcap E | A \in \mathcal{A}\}$ now as $E \in \mathcal{A}$ we have

$$\begin{array}{lcl} A \in \mathcal{A}_E & \Rightarrow & \exists A' \in \mathcal{A} \text{ so that } A = A' \bigcap E \\ & \underset{A' \bigcap E \in \mathcal{A} \text{ and } A' \bigcap E \subseteq E}{\Rightarrow} & A' \in \{A \in \mathcal{A} | A \subseteq E\} \\ A \in \{A \in \mathcal{A} | A \subseteq E\} & \Rightarrow & A \in \mathcal{A} \wedge A \subseteq E \\ & \Rightarrow & A \in \mathcal{A} \wedge A = A \bigcap E \\ & \Rightarrow & A \in A_E \end{array}$$

   which proves that $\mathcal{A}_E = \{A \in \mathcal{A} | A \subseteq E\} \subseteq \mathcal{A}$

2. Let $A \in \mathcal{A}_E$. Using the definition of μ_E [see 18.36] we have that $\mu_E(A) = \inf(\{\mu(B) | B \in \mathcal{A} \wedge A \subseteq B\})$. Take $A \in \mathcal{A}_E \subseteq_{(1)} \mathcal{A}$. Then $\forall B \in \mathcal{A}$ with $A \subseteq B$ we have by 18.20 that $\mu_{|E}(A) = \mu(A) \leqslant \mu(B)$, hence $\mu_E(A) \leqslant \inf(\{\mu(B) | B \in \mathcal{A} \wedge E \subseteq B\})$ proving that

$$\mu_{|\mathcal{A}_E}(A) \leqslant \mu_E(A) \tag{18.339}$$

   Further as $A \in \mathcal{A}$ we have, as trivially $A \subseteq A$, that $\mu(A) \in \{\mu(B) | B \in \mathcal{A} \wedge E \subseteq B\}$ and thus that $\inf(\mu(B) | B \in \mathcal{A} \wedge E \subseteq B) \leqslant \mu(A)$ or $\mu_E(A) \leqslant \mu(A) = \mu_{|\mathcal{A}_E}(A)$. This together with (18.339) proves

$$\forall A \in \mathcal{A}_E \text{ we have } \mu_E(A) = \mu_{|\mathcal{A}_E}(A) \text{ or } \mu_E = \mu_{|A_E}$$

   proving (2). □

As we have to deal with negligible sets, conegligible sets and measurable sets the following lemma is important.

Lemma 18.153. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space, $E \in \mathcal{A}$ then for the subspace $\langle E, \mathcal{A}_E, \mu_E\rangle$ we have*

1. *$\forall H \subseteq E$ we have that H is μ_E-negligible if and only if H is μ-negligible*
2. *If $H \subseteq X$ is μ-conegligible then $H \bigcap E$ is μ_E-conegligible.*

3. *If $F \in \mathcal{A}_E$ then for the subspace $\langle F, (\mathcal{A}_E)_{|F}, \mu_F \rangle$ we have that*
   a. $(\mathcal{A}_E)_F = \mathcal{A}_F$
   b. $(\mu_E)_F = \mu_F$
4. *If $H \subseteq E$ is μ_E-conegligible then $H \bigcup (X \setminus E)$ is μ-conegligible.*

Proof.

1. Let $H \subseteq E$. If H is μ_E-negligible there exists a $A \in \mathcal{A}_E$ with $\mu_E(A) = 0$ and $H \subseteq A$. As $\mu_E(A) \underset{18.152\ (2)}{=} \mu_{|\mathcal{A}_E}(A) = \mu(A)$ we have that $\mu(A) = 0$ which as $A \in \mathcal{A}$ [see 18.152 (1)] it follows that A is μ-negligible. Next if H is μ-negligible then there exists a $A \in \mathcal{A}$ with $\mu(A) = 0$ such that $H \subseteq A$. So $H \underset{H \subseteq E}{=} H \bigcap E \subseteq A \bigcap E \in \mathcal{A}_E$ [see 18.36] and $\mu_{|E}(A \bigcap E) \underset{18.152}{=} \mu_{|\mathcal{A}_E}(A \bigcap E) = \mu(A \bigcap E) \subseteq \mu(E) = 0$ which proves that H is $\mu_{|E}$-negligible.
2. As H is μ-conegligible there exists a set $F \in \mathcal{A}$ with $\mu(F) = 0$ such that $X \setminus H \subseteq F$. Now

$$\begin{array}{rcl} E \setminus \left(E \bigcap H\right) & \underset{1.31}{=} & (E \setminus E) \bigcup (E \setminus H) \\ & = & E \setminus H \\ & \underset{E \subseteq X}{=} & \left(X \bigcap E\right) \setminus H \\ & \underset{1.31(10)}{=} & E \bigcap (X \setminus H) \\ & \subseteq & E \bigcap F \end{array}$$

   So as $E \bigcap F \in \mathcal{A}_E$ and $\mu_E(E \bigcap F) = \mu(E \bigcap F) \leqslant \mu(F) = 0$ we have that $E \setminus (E \bigcap H)$ is μ_E-negligible proving that

$$E \bigcap H \text{ is } \mu_E\text{-conegligible}$$

3. We have
   a. As $F \in \mathcal{A}_E$ we have that $F \in \mathcal{A}$ and $F \subseteq E$. If now $A \in (\mathcal{A}_E)_F$ then $A \in \mathcal{A}_E \subseteq \mathcal{A}$ [see 18.152] and $A \subseteq F$ hence $A \in \mathcal{A}_F$. If $A \in \mathcal{A}_F$ then $A \in \mathcal{A}$ and $A \subseteq F$, as $F \subseteq E$ we have $A \in \mathcal{A}_E$, further $A \subseteq F$ proving that $A \in (\mathcal{A}_E)_F$. So we have that

$$(\mathcal{A}_E)_F = \mathcal{A}_F$$

   b. $(\mu_E)_F \underset{18.152}{=} (\mu_E)_{|(\mathcal{A}_E)_F} \underset{18.152}{=} (\mu_{|\mathcal{A}_E})_{|(\mathcal{A}_E)_F} \underset{2.28}{=} \mu_{|\mathcal{A}_E \bigcap (\mathcal{A}_E)_F} \underset{(\mathcal{A}_E)_F \subseteq \mathcal{A}_E}{=} \mu_{|(\mathcal{A}_E)_F} \underset{(a)}{=} \mu_F$
4. As H is μ_E-conegligible there exists a $A \in \mathcal{A}_E$ with $\mu_E(A) = 0$ such that $E \setminus H \subseteq A$. Now $X \setminus (H \bigcup (X \setminus E)) \underset{1.31}{=} (X \setminus (X \setminus E)) \setminus H = E \setminus H \subseteq A$ and as $A \in \mathcal{A}_E \subseteq \mathcal{A}$ we have that $\mu(A) = \mu_{|\mathcal{A}_E}(A) = \mu_E(A) = 0$. So we have that $X \setminus (H \bigcup (X \setminus E))$ is μ-negligible or $H \bigcup (X \setminus E)$ is μ-conegligible. □

Lemma 18.154. *Let $\langle X, \mathcal{A}, \mu \rangle$ be a measure space, $E \in \mathcal{A}$ then we have*

1. *If $f: X \to \mathbb{R}$ is a partial function then $\forall a \in \mathbb{R}$ we have if $a \leqslant_{a.e.[\mu]} f$ then $a \leqslant_{a.e.[\mu_E]} f_{|E}$*
2. *If $f, g: X \to \mathbb{R}$ are partial functions such that $f \leqslant_{a.e.[\mu]} g$ then $f_{|E} \leqslant g_{|E}$*

3. *If $f, g: X \rightarrow \mathbb{R}$ are partial functions such that $f =_{a.e.[\mu]} g$ then $f_{|E} = g_{|E}$*

Proof.

1. As $a \leqslant_{a.e.[\mu]} f$ we have that $\{a \leqslant f\}$ is μ-conegligible, hence $\{a \leqslant f\} \bigcap E$ is by the previous lemma [see 18.153] μ_E-conegligible. Further we have

$$\begin{aligned} x \in \{a \leqslant f\} \bigcap E &\Leftrightarrow x \in \operatorname{dom}(f) \bigcap E \wedge a \leqslant f(x) \\ &\Leftrightarrow x \in \operatorname{dom}(f) \bigcap E \wedge a \leqslant f_{|E}(x) \\ &\Leftrightarrow x \in \operatorname{dom}(f_{|E}) \wedge a \leqslant f_{|E}(x) \\ &\Leftrightarrow x \in \{a \leqslant f_{|E}\} \end{aligned}$$

   proving that $\{a \leqslant f_{|E}\}$ is μ_E-conegligible. Hence $a \leqslant_{a.e.[\mu]} f$

2. As $f \leqslant_{a.e.[\mu]} g$ we have that $\{x \in \operatorname{dom}(f) \bigcap \operatorname{dom}(g) | f(x) \leqslant g(x)\}$ is μ-conegligible. Using the previous lemma [see 18.153] it follows then that $\{x \in \operatorname{dom}(f) \bigcap \operatorname{dom}(g) | f(x) \leqslant g(x)\} \bigcap E$ is μ_E-conegligible. Further we have

$$\begin{aligned} x \in \big\{x \in \operatorname{dom}(f) \bigcap \operatorname{dom}(g) | f(x) \leqslant g(x)\big\} \bigcap E &\Leftrightarrow x \in \operatorname{dom}(f) \bigcap \operatorname{dom}(g) \bigcap E \wedge f(x) \leqslant g(x) \\ &\Leftrightarrow x \in \operatorname{dom}(f) \bigcap \operatorname{dom}(g) \bigcap E \wedge f_{|E}(x) \leqslant g_{|E}(x) \\ &\Leftrightarrow x \in \operatorname{dom}(f_{|E}) \bigcap \operatorname{dom}(g_{|E}) \wedge f_{|E}(x) \leqslant g_{|E}(x) \\ &\Leftrightarrow x \in \big\{x \in \operatorname{dom}(f_{|E}) \bigcap \operatorname{dom}(g_{|E}) | f_{|E}(x) \leqslant g_{|E}(x)\big\} \end{aligned}$$

   proving that $\{x \in \operatorname{dom}(f_{|E}) \bigcap \operatorname{dom}(g_{|E}) | f_{|E}(x) \leqslant g_{|E}(x)\}$ is μ_E-conegligible. Hence $f_{|E} \leqslant_{a.e.[\mu_E]} g_{|E}$.

3. As $f =_{a.e.[\mu]} g$ we have by 18.30 that $f \leqslant_{a.e.[\mu]} g$ and $g \leqslant_{a.e.[\mu]} f$ hence using (2) we have $f_{|E} \leqslant_{a.e.[\mu_E]} g_{|E}$ and $g_{|E} \leqslant_{a.e.[\mu_E]} f_{|E}$. Applying 18.30 again proves $f_{|E} =_{a.e.[\mu_E]} g_{|E}$ □

Definition 18.155. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space, $E \in \mathcal{A}$ and $f: X \rightarrow \mathbb{R}$ a partial function then if $f_{|E} \in \mathcal{L}[E, \mathcal{A}_E, \mu_E]$ we define $\int_E f d\mu$ as $\int_E f d\mu = \int f_{|E} d\mu_E$*

Note 18.156. If $E = X$ then of course $\langle X, \mathcal{A}, \mu\rangle = \langle X, \mathcal{A}_X, \mu_X\rangle$ and $\int_X f d\mu = \int f d\mu$

Definition 18.157. *Let X be a set $E \subseteq X$ and $f: E \rightarrow \mathbb{R}$ a partial function [so $\operatorname{dom}(f) \subseteq E$] then $f^E: X \rightarrow \mathbb{R}$ is the partial function defined by $f^E = \begin{cases} 0 \text{ if } x \in X \setminus E \\ f(x) \text{ if } x \in \operatorname{dom}(f) \end{cases}$ [which is well defined because $(X \setminus E) \bigcap \operatorname{dom}(f) = \emptyset$]. Clearly $\operatorname{dom}(f^E) = \operatorname{dom}(f) \bigsqcup (X \setminus A)$*

Lemma 18.158. *Let X be a set $E \subseteq X$ and $f: E \rightarrow \mathbb{R}$ a partial function then $|f|^E = |f^E|$*

Proof. Let $x\in\mathrm{dom}(|f|)\bigcup(X\setminus E)\underset{\mathrm{dom}(f)=\mathrm{dom}(|f|)}{=}\mathrm{dom}(f)\bigcup(X\setminus E)$ then we have either

$\boldsymbol{x\in X\setminus E}$**.** then $|f|^E(x)=0=|0|=|f^E(x)|=|f^E|(x)$

$\boldsymbol{x\in \mathrm{dom}(f)}$**.** then $|f|^E(x)=|f|(x)=|f(x)|=|f^E(x)|=|f^E|(x)$ □

Lemma 18.159. *Let $\langle X,\mathcal{A},\mu\rangle$ be a measure space, $E\in\mathcal{A}$ and $f{:}E\rightarrow\mathbb{R}$, $g{:}E\rightarrow\mathbb{R}$ be partial functions then we have*

1. *If $g\leqslant_{a.e.[\mu_E]}f$ then $g^E\leqslant_{a.e.[\mu]}f^E$*
2. *If $g=_{a.e.[\mu_E]}f$ then $g^E=_{a.e.[\mu]}f^E$*

Proof.

1. By assumption $F=\{x\in\mathrm{dom}(g)\bigcap\mathrm{dom}(f)|g(x)\leqslant f(x)\}$ is μ_E-conegligible. Take now $x\in(X\setminus E)\bigcup F\subseteq(X\setminus E)\bigcup(\mathrm{dom}(f)\bigcap\mathrm{dom}(g))=\mathrm{dom}(g^E)\bigcap\mathrm{dom}(f^E)$ then we have [as $F\subseteq\mathrm{dom}(f)\bigcap\mathrm{dom}(g)\subseteq E$] either

   $\boldsymbol{x\in X\setminus E}$**.** then $g^E(x)=0\leqslant 0=f^E(x)$

   $\boldsymbol{x\in F}$**.** then as $F\subseteq\mathrm{dom}(f)\bigcap\mathrm{dom}(g)$ we have $g^E(x)=g(x)\leqslant f(x)=f^E(x)$

   proving that $\forall x\in(X\setminus E)\bigcup F\;\; g^E(x)\leqslant f^E(x)$. So as by 18.153 (4) $(X\setminus E)\bigcup F$ is μ-conegligible it follows that

   $$g^E\leqslant_{a.e.[\mu]}f^E$$

2. If $g=_{a.e.[\mu]}f$ then by 18.30 we have $g\leqslant_{a.e.[\mu]}f$ and $f\leqslant_{a.e.[\mu]}g$ so that by (1) we have $g^E\leqslant_{a.e.[\mu]}f^E$and $f^E\leqslant_{a.e.[\mu]}g^E$. Using 18.30 again proves that $g=_{a.e.[\mu]}f$. □

In this section we are dealing with partial functions on a subset of a given set, this apply also to characteristics functions [see 2.14]. So we need a notation to specify the set where the characteristics function is defined on.

Notation 18.160. *Let X be a set and $E\subseteq X$ then for every $A\subseteq A$ we note the characteristics **function** defined by A on E by $\mathcal{X}_{E,A}$. So $\mathcal{X}_{E,A}{:}X\rightarrow\mathbb{R}$ is defined by $\mathcal{X}_{E,A}=\left\{\begin{smallmatrix}1\text{ if }x\in A\\ 0\text{ if }x\in E\setminus A\end{smallmatrix}\right.$*

Lemma 18.161. *Let X be a set, $E\subseteq X$ and $f{:}E\rightarrow\mathbb{R}$ the **function** defined by $f=\sum_{i=1}^n\alpha_i\cdot\mathcal{X}_{E,A_i}$ where $\forall i\in\{1,\ldots,n\}$ $A_i\subseteq E$ then $f^E{:}X\rightarrow\mathbb{R}$ is the **function** defined by $f^E=\sum_{i=1}^n\alpha_i\cdot\mathcal{X}_{A_i}$*

Proof. First take $A\subseteq E$ then $\forall x\in E$ we have either

$\boldsymbol{x\in E\setminus A}$**.** then $\mathcal{X}_{E,A}(x)=0=\mathcal{X}_A(x)$

$\boldsymbol{x\in A}$**.** then $\mathcal{X}_{E,A}(x)=1=\mathcal{X}_A(x)$

proving that

$$\forall A\subseteq E\text{ we have }\forall x\in E\text{ that }\mathcal{X}_{E,A}(x)=\mathcal{X}_A(x)\tag{18.340}$$

Second as $\operatorname{dom}(f)=E$ we have that $\operatorname{dom}(f^E)=\operatorname{dom}(f)\bigcup(X\setminus E)=E\bigcup(X\setminus E)=X$ so that $f^E: X\rightarrow\mathbb{R}$ is a function. Further if $x\in X$ then we have either

$\boldsymbol{x\in X\setminus E.}$ then $f^E(x)=0$ and as $\forall i\in\{1,\ldots,n\}$ $A_i\subseteq E$ it follows that $x\in X\setminus A_i$. Hence $(\sum_{i=1}^n\alpha_i\cdot\mathcal{X}_{A_i})(x)=\sum_{i=1}^n\alpha_i\cdot\mathcal{X}_{A_i}(x)=0=f^E(x)$

$\boldsymbol{x\in E=\operatorname{dom}(f).}$ the we have $(\sum_{i=1}^n\alpha_i\cdot\mathcal{X}_{A_i})(x)=\sum_{i=1}^n\alpha_i\cdot\mathcal{X}_{A_i}(x)\underset{(18.340)}{=}\sum_{i=1}^n\alpha_i\cdot\mathcal{X}_{E,A_i}(x)=f(x)=f^E(x)$

proving that

$$f^E=\sum_{i=1}^n\alpha_i\cdot\mathcal{X}_{A_i}\qquad\square$$

Lemma 18.162. *Let $\langle X,\mathcal{A},\mu\rangle$ be a measure space, $E\in\mathcal{A}$ and $f\in\mathcal{S}[E,\mathcal{A}_E]$ then $f^E\in\mathcal{S}[X,\mathcal{A}]$ and $\int^S fd\mu_E=\int^S f^E d\mu$ [see 18.140] we have $\int fd\mu_E=\int f^E d\mu$]*

Proof. Let $f\in\mathcal{S}[E,\mathcal{A}_E]$ then there exists a representation $f=\sum_{i=1}^n\alpha_i\cdot\mathcal{X}_{E,A_i}$ where $\forall i\in\{1,\ldots,n\}$ $A_i\in\mathcal{A}_E$, $\mu_E(A_i)<\infty$. Let $i\in\{1,\ldots,n\}$ then by 18.152 $A_i\in\mathcal{A}$ and $\mu(A_i)=\mu_E(A_i)<\infty$. As $f^E\underset{18.161}{=}\sum_{i=1}^n\alpha_i\cdot\mathcal{X}_{A_i}$ we have that $f^E\in\mathcal{S}[X,\mathcal{A}]$. Further

$$\begin{aligned}\int^S fd\mu_E &= \sum_{i=1}^n\alpha_i\cdot\mu_E(A_i)\\ &\underset{18.152}{=} \sum_{i=1}^n\alpha_i\cdot\mu_{|\mathcal{A}_E}(A_i)\\ &\underset{\forall i\in\{1,\ldots,n\}A_i\in\mathcal{A}_E}{=} \sum_{i=1}^n\alpha_i\cdot\mu(A_i)\\ &= \int^S f^E d\mu\end{aligned}$$

$\square$

Lemma 18.163. *Let $\langle X,\mathcal{A},\mu\rangle$ be a measure space, $E\in\mathcal{A}$ and $f\in\mathcal{L}_+[E,\mathcal{A}_E,\mu_E]$ then $f^E\in\mathcal{L}_+[X,\mathcal{A},\mu]$ and $\int^+ fd\mu_E=\int^+ f^E d\mu$ [so using 18.140 we have $\int fd\mu_E=\int f^E d\mu$]*

Proof. By definition [see Definition 18.126] we have that $0\leqslant f$ and there exists a $\{f_n\}_{n\in\mathbb{N}}\subseteq\mathcal{S}[E,\mathcal{A}_E]$ such that

$$\forall n\in\mathbb{N}\vDash 0\leqslant f_n\leqslant f_{n+1},\ f=_{a.e.[\mu_E]}\lim_{n\rightarrow\infty}f_n\ \text{ and }\ \sup\left(\left\{\int^S f_n d\mu_E|n\in\mathbb{N}\right\}\right)<\infty\tag{18.341}$$

Using 18.162 we have

$$\{(f_n)^E\}_{n\in\mathbb{N}}\subseteq\mathcal{S}[X,\mathcal{A}]\wedge\forall n\in\mathbb{N}\vDash\int^S f_n d\mu_E=\int^S(f_n)^E d\mu\tag{18.342}$$

Using the above it follows that

$$\sup\left(\left\{\int^S(f_n)^E d\mu_E|n\in\mathbb{N}\right\}\right)=\sup\left(\left\{\int^S f_n d\mu|n\in\mathbb{N}\right\}\right)<\infty\tag{18.343}$$

Further as $\forall n\in\mathbb{N},\forall x\in X$ we have $0\leqslant\left\{\begin{smallmatrix}0 \text{ if } x\in X\setminus E\\ f_n(x) \text{ if } x\in E\end{smallmatrix}\right.\leqslant\left\{\begin{smallmatrix}0 \text{ if } x\in X\setminus E\\ f_{n+1}(x) \text{ if } x\in E\end{smallmatrix}\right.$ proving that

$$\forall n\in\mathbb{N} \text{ we have } 0\leqslant(f_n)^E\leqslant(f_{n+1})^E \tag{18.344}$$

Now if $x\in\mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right)\bigcup X\setminus E$ then we have either [as $\mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right)\subseteq E$ (see 18.124)]]

$x\in X\setminus E$. then $\forall n\in\mathbb{N}$ we have $(f_n)^E(x)=0$ such that $\lim_{n\rightarrow\infty}(f_n)^E(x)$ exists and is equal to 0. So $\lim_{n\rightarrow\infty}(f_n)^E(x)=0=\left(\lim_{n\rightarrow\infty} f_n\right)^E(x)$

$x\in\mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right)$. then as $f_n^E(x)=f_n(x)$ $\lim_{n\rightarrow\infty} f_n^E(x)$ exists and $\lim_{n\rightarrow\infty} f_n^E(x)=\lim_{n\rightarrow\infty} f_n(x)=\left(\lim_{n\rightarrow\infty} f_n\right)^E(x)$

which proves that

$$\mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right)\bigcup (X\setminus E)\subseteq\mathrm{dom}\left(\lim_{n\rightarrow\infty}(f_n)^E\right) \tag{18.345}$$

and

$$\forall x\in\mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right)\bigcup (X\setminus E)\vDash\left(\lim_{n\rightarrow\infty}(f_n)^E\right)(x)=\left(\lim_{n\rightarrow\infty} f_n\right)^E(x) \tag{18.346}$$

Further if $x\in\mathrm{dom}\left(\lim_{n\rightarrow\infty}(f_n)^E\right)\underset{18.124}{=}\left\{x\in X|\lim_{n\rightarrow\infty}(f_n)^E \text{exists}\right\}$ then we have either

$x\in X\setminus E$. then $x\in\mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right)\bigcup (X\setminus E)$

$x\in E$. then as $x\in\mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n^E\right)$ $\lim_{n\rightarrow\infty} f_n^E(x)$ exists, further $\forall n\in\mathbb{N}$ $\mathrm{dom}(f_n)=E$ so that $f_n(x)=f_n^E(x)$, hence it follows that $\lim_{n\rightarrow\infty} f_n(x)$ exists proving that $x\in\mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right)\subseteq\mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right)\bigcup (X\setminus E)$

which proves that $\mathrm{dom}\left(\lim_{n\rightarrow\infty}(f_n)^E\right)\subseteq\mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right)\bigcup (X\setminus E)$. Combining this with (18.345) proves that $\mathrm{dom}\left(\lim_{n\rightarrow\infty}(f_n)^E\right)=\mathrm{dom}\left(\lim_{n\rightarrow\infty} f_n\right)\bigcup (X\setminus E)$, finally applying (18.346) proves

$$\left(\lim_{n\rightarrow\infty} f_n\right)^E=\lim_{n\rightarrow\infty}(f_n)^E \tag{18.347}$$

As $f=_{a.e.[\mu_E]}\lim_{n\rightarrow\infty} f_n$ we have by (18.159) that $f^E=_{a.e.[\mu]}\left(\lim_{n\rightarrow\infty} f_n\right)^E\underset{(18.347)}{=}\lim_{n\rightarrow\infty}(f_n)^E$ giving

$$f^E\underset{a.e.[\mu]}{=}\lim_{n\rightarrow\infty}(f_n)^E \tag{18.348}$$

The conditions for a positive integrable function [see Definition 18.126] are then satisfied by (18.342),(18.343),(18.344) and (18.348) so that we have

$$f^E\in\mathcal{L}_+[X.\mathcal{A}.\mu] \tag{18.349}$$

Using (18.344) and (18.349) allows us to apply proposition 18.132 giving $\int^{+} f^E d\mu = \sup\left(\left\{\int^S (f_n)^E d\mu | n \in \mathbb{N}\right\}\right) \underset{(18.343)}{=} \sup\left(\left\{\int^S f_n d\mu_E | n \in \mathbb{N}\right\}\right) \underset{18.132}{=} \int^{+} f d\mu_E$ proving that

$$\int^{+} f d\mu_E = \int^{+} f^E d\mu \tag{18.350}$$ □

Proposition 18.164. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space, $E \in \mathcal{A}$ and $f: E \rightarrow \mathbb{R}$ a partial function then for $f^E: X \rightarrow \mathbb{R}$ we have that $f \in \mathcal{L}[E, \mathcal{A}_E, \mu_E]$ if and only if $f^E \in \mathcal{L}[X, \mathcal{A}, \mu]$. If $f \in \mathcal{L}[E, \mathcal{A}_E, \mu_E]$ [or $f^E \in \mathcal{L}[X, \mathcal{A}, \mu]$] then $\int f d\mu_E = \int f^E d\mu$.*

Proof. Let $f \in \mathcal{L}[E, \mathcal{A}_E, \mu_E]$ then $f = f_1 - f_2$ where $f_1, f_2 \in \mathcal{L}_+[E, \mathcal{A}_E]$ and $\int f d\mu_E = \int^{+} f_1 d\mu_E - \int^{+} f_2 d\mu_E$. Using 18.163 we have $(f_1)^E, (f_2)^E \in \mathcal{L}[X, \mathcal{A}, \mu]$ and $\int^{+} f_1 d\mu_E = \int^{+} (f_1)^E d\mu$, $\int^{+} f_2 d\mu_E = \int^{+} (f_2)^E d\mu$. Now for $x \in \mathrm{dom}(f^E) = \mathrm{dom}(f_1 - f_2) \bigcup (X \setminus E) = (\mathrm{dom}(f_1) \bigcap \mathrm{dom}(f_2)) \bigcup (X \setminus E) = \mathrm{dom}((f_1)^E) \bigcap \mathrm{dom}((f_2)^E)$ we have

$$\begin{aligned}
f^E(x) &= \begin{cases} 0 \text{ if } x \in X \setminus E \\ f(x) \text{ if } x \in \mathrm{dom}(f) \end{cases} \\
&= \begin{cases} 0 \text{ if } x \in X \setminus E \\ f_1(x) \text{ if } x \in \mathrm{dom}(f) \end{cases} - \begin{cases} 0 \text{ if } x \in X \setminus E \\ f_2(x) \text{ if } x \in \mathrm{dom}(f) \end{cases} \\
&\underset{\mathrm{dom}(f) \subseteq \mathrm{dom}(f_1), \mathrm{dom}(f_2)}{=} (f_1)^E(x) - (f_2)^E(x)
\end{aligned}$$

proving that $f^E = (f_1)^E - (f_2)^E$. Then by definition we have as $(f_1)^E, (f_2)^E \in \mathcal{L}[X, \mathcal{A}, \mu]$ that

$$f^E \in \mathcal{L}[X, \mathcal{A}, \mu] \tag{18.351}$$

and $\int f^E d\mu = \int^{+} (f_1)^E d\mu - \int^{+} (f_2)^E d\mu = \int^{+} f_1 d\mu_E - \int^{+} f_2 d\mu_E = \int f d\mu_E$ giving

$$\int f^E d\mu = \int f d\mu_E \tag{18.352}$$

For the opposite implication. Let $f^E \in \mathcal{L}[X, \mathcal{A}, \mu]$. Using 18.143 there exists a μ-conegligible set $F \subseteq X$ such that $(f^E)_F$ is $\mathcal{A}$-measurable, further by 18.137 $\mathrm{dom}(f^E)$ is μ-conegligible hence there exists by 18.27 a μ-conegligible set $G \in \mathcal{A}$ such that $G \subseteq F \bigcap \mathrm{dom}(f^E)$. As $(f^E)_{|F}$ is $\mathcal{A}$-measurable we have by 18.103 (7) that $((f^E)_{|F})_{|G}$ is $\mathcal{A}$-measurable which as $((f^E)_{|F})_{|G} \underset{G \subseteq F \text{ and } 2.28}{=} (f^E)_{|G}$ proves that $(f^E)_{|G}$ is $\mathcal{A}$-measurable. Summarizing we have

$$\exists G \in \mathcal{A} \text{ such that } G \text{ is } \mu\text{-conegligible, } G \subseteq \mathrm{dom}(f^E) \text{ and } (f^E)_{|G} \text{ is } \mathcal{A}\text{-measurable} \tag{18.353}$$

Using 18.153(2) we have that $G \bigcap E$ is μ_E-conegligible. Further $G \bigcap E \subseteq \mathrm{dom}(f^E) \bigcap E = (\mathrm{dom}(f) \bigcup (X \setminus E)) \bigcap E = \mathrm{dom}(f) \bigcap E \underset{\mathrm{dom}(f) \subseteq E}{=} \mathrm{dom}(f) = \mathrm{dom}(|f|)$. So we have

$$E \bigcap G \text{ is } \mu_E\text{-conegligible, } E \bigcap G \subseteq \mathrm{dom}(f) = \mathrm{dom}(|f|) \text{ and as } G \in \mathcal{A} \text{ we have } E \bigcap G \in \mathcal{A} \tag{18.354}$$

Next let $a \in \mathbb{R}$. If $x \in \{a \leqslant f_{|E \bigcap G}\}$ we have that $x \in \operatorname{dom}(f) \bigcap (E \bigcap G)$ and $a \leqslant f_{|E \bigcap G}(x) = f(x) \underset{x \in \operatorname{dom}(f)}{=} f^E(x) \underset{x \in G}{=} (f^E)_{|G}(x)$ proving that $x \in E \bigcap \{a \leqslant (f^E)_{|G}\}$. Further if $x \in E \bigcap \{a \leqslant (f^E)_{|G}\}$ then $x \in E \bigcap \operatorname{dom}((f^E)_{|G}) = E \bigcap G \bigcap (\operatorname{dom}(f) \bigcup (X \setminus E)) = E \bigcap G \bigcap \operatorname{dom}(f)$ and $a \leqslant (f^E)_{|G}(x) = (f^E)(x) \underset{x \in \operatorname{dom}(f)}{=} f(x) \underset{x \in \operatorname{dom}(E \bigcap G)}{=} f_{|E \bigcap G}(x)$ proving that $x \in \{a \leqslant f_{|E \bigcap G}\}$. So we have that $\{a \leqslant f_{|E \bigcap G}\} = E \bigcap \{a \leqslant (f^E)_{|G}\} \in \mathcal{A}_E$ [as $(f^E)_{|G}$ is $\mathcal{A}$-measurable so that $\{a \leqslant (f^E)_{|G}\} \in \mathcal{A}$]. Hence $f_{|E \bigcap G}$ is $\mathcal{A}_E$-measurable and applying then 18.103 (8) proves that $|f_{|E \bigcap G}|$ is $\mathcal{A}_E$-measurable. As $|f_{|E \bigcap G}| \underset{18.90}{=} |f|_{|E \bigcap G}$ we have

$$|f|_{|E \bigcap G} \text{ is } \mathcal{A}_E\text{-measurable} \tag{18.355}$$

If $x \in \{a \leqslant |f|_{|E \bigcap G}\}$ then $x \in \operatorname{dom}(|f|_{|E \bigcap G}) = \operatorname{dom}(|f|) \bigcap E \bigcap G = \operatorname{dom}(f) \bigcap E \bigcap G$ and $a \leqslant (|f|_{|E \bigcap G})(x) = |f(x)| \underset{x \in \operatorname{dom}(f)}{=} |f^E(x)|$ proving that $x \in \{a \leqslant |f^E|\}$. Hence

$$\forall a \in \mathbb{R} \text{ we have } \{a \leqslant |f|_{|E \bigcap G}\} \subseteq \{a \leqslant |f^E|\} \tag{18.356}$$

As by (18.355) $|f|_{|E \bigcap G}$ is $\mathcal{A}_E$-measurable we have that $\forall a \in \mathbb{R}\ \{a \leqslant |f|_{|E \bigcap G}\} \in \mathcal{A}_E \subseteq \mathcal{A}$ so that $\mu(\{a \leqslant |f|_{|E \bigcap G}\})$ exists. Also as $f^E \in \mathcal{L}[X, \mathcal{A}, \mu]$ it follows from 18.143 (3) that $|f^E| \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ and using 18.130 we have that $\forall \varepsilon > 0\ \mu(\{\varepsilon \leqslant |f^E|\}) < \infty$. So

$$\forall \varepsilon > 0 \text{ we have } \mu_E(\{\varepsilon \leqslant |f|_{|E \bigcap G}\}) = \mu(\{\varepsilon \leqslant |f|_{|E \bigcap G}\}) \leqslant_{(18.356) \text{ and } 18.20} \mu(\{\varepsilon \leqslant |f^E|\}) < \infty \tag{18.357}$$

Let $g \in \mathcal{S}[E, \mathcal{A}_E]$ such that $g \leqslant_{a.e.[\mu_E]} |f|$ then using (18.159) $g^E \leqslant_{a.e.[\mu]} |f|^E \underset{(18.158)}{=} |f^E|$. So we have that $\int^S g\, d\mu_E \underset{18.162}{=} \int g^E d\mu \leqslant_{18.141\ (4)} \int |f^E| d\mu$, as $g \in \mathcal{S}[E, \mathcal{A}_E]$ was chosen arbitrary we have then

$$\sup\left(\left\{\int^S g\, d\mu_E | g \in \mathcal{S}[E, \mathcal{A}_E] \wedge g \leqslant_{a.e.[\mu_E]} |f|\right\}\right) \leqslant \int |f^E| d\mu < \infty \tag{18.358}$$

Now (18.354),(18.355) and (18.358) allows us to apply 18.130 (1) proving that $|f| \in \mathcal{L}_+[E, \mathcal{A}_E, \mu_E]$. Finally using 18.143(3) together with (18.355) and (18.354) proves that $f \in [E, \mathcal{A}_E, \mu_E]$ so we have

$$\text{If } f^E \in \mathcal{L}[E, \mathcal{A}_E, \mu_E] \text{ then } f \in \mathcal{L}[E, \mathcal{A}, \mu_E] \tag{18.359}$$

The proposition is then proved by (18.351),(18.352) and (18.359). □

Corollary 18.165. *Let $\langle X, \mathcal{A}, \mu \rangle$ be a measure space, $E \in \mathcal{A}$ and $f: X \to \mathbb{R}$ a partial function such that $f \in \mathcal{L}[X, \mathcal{A}, \mu]$ then $f_{|E} \in \mathcal{L}[\mathcal{A}, \mathcal{A}_E, \mu]$.*

Proof. Then $\operatorname{dom}(f)$ is μ-conegligible $\{f \in \mathcal{L}[X, \mathcal{A}, \mu]$ together with 18.137]. Consider $(f_{|E})^E$ then $\operatorname{dom}((f_{|E})^E) = (X \setminus E) \bigcup \operatorname{dom}(f_{|E}) = (X \setminus E) \bigcup (\operatorname{dom}(f) \bigcap E) = ((X \setminus E) \bigcup \operatorname{dom}(f)) \bigcap ((X \setminus E) \bigcup E) = ((X \setminus E) \bigcup \operatorname{dom}(f)) \bigcap X = (X \setminus E) \bigcup \operatorname{dom}(f)$ which by 18.27 (2) proves that $\operatorname{dom}((f_{|E})^E)$ is μ-conegligible. Further $\forall x \in \operatorname{dom}(f)$ we have either

- $\boldsymbol{x \in \operatorname{dom}(f) \setminus E}$**.** then $x \in X \setminus E$ so that $|(f_{|E})^E|(x) = 0 \leqslant |f(x)| = |f|(x)$
- $\boldsymbol{x \in \operatorname{dom}(f) \bigcap E = \operatorname{dom}(f_{|E})}$**.** then $|(f_{|E})^E|(x) = |(f_{|E})^E(x)| = |f_{|E}(x)| = |f(x)| = |f|(x)$

so $\forall x \in \text{dom}(f)$ we have $|(f_{|E})^E|(x) \leqslant |f|(x)$ so that $|(f_{|E})^E| \leqslant a.e.[\mu] |f|$. As $(f_{|E})_|$ $\square$

We proceed now to prove that every integrable function is integrable in a subspace. First we prove this by positive integrable functions. We prove this in phases, first for simple functions, then positive integrable functions and finally integrable functions.

Lemma 18.166. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space, $E \in \mathcal{A}$ and $f \in \mathcal{S}[X, \mathcal{A}]$ with representation $f = \sum_{i=1}^n \alpha_i \cdot \mathcal{X}_{A_i}$ then $f_{|E} \in \mathcal{S}[E, \mathcal{A}_E]$ and has representation $\sum_{i=1}^n \alpha_i \cdot \mathcal{X}_{E, A_i \bigcap E}$. Further $\int^S f_{|E} d\mu_E \leqslant \int^S f d\mu$.*

Proof. Let $A \in \mathcal{A}$ then we have for $x \in E$ either

$x \in E \setminus A$. then $(\mathcal{X}_A)_{|E}(x) \underset{x \in E}{=} \mathcal{X}_A(x) \underset{x \notin A}{=} 0 \underset{x \notin A \bigcap E}{=} \mathcal{X}_{E, A \bigcap E}(x)$

$x \in A$. then $(\mathcal{X}_A)_{|E}(x) \underset{x \in E}{=} \mathcal{X}_A(x) \underset{x \in A}{=} 1 \underset{x \in A \bigcap E}{=} \mathcal{X}_{E, A \bigcap E}(x)$

proving that

$$\forall A \in \mathcal{A} \text{ we have } (\mathcal{X}_A)_{|E} = \mathcal{X}_{E, A \bigcap E} \tag{18.360}$$

Take now $f \in \mathcal{S}[X, \mathcal{A}]$ with representation $\sum_{i=1}^n \alpha_i \cdot \mathcal{X}_{A_i}$ then $\{A_i\}_{i \in \{1,\ldots,n\}} \subseteq \mathcal{A}$ and $\forall i \in \{1,\ldots,n\}$ $\mu(A_i) < \infty$. So using 18.153 we have that $\{A_i \bigcap E\}_{i \in \{1,\ldots,n\}} \subseteq \mathcal{A}_E$, also $\forall i \in \{1,\ldots,n\}$ we have $\mu_E(A \bigcap E_i) \underset{18.152}{=} \mu_{|\mathcal{A}_E}(A_i) = \mu(A_i \bigcap E) \leqslant \mu(A_i) < \infty$. Further for $x \in E$ we have $f_{|E}(x) = f(x) = \sum_{i=1}^n \alpha_i \cdot \mathcal{X}_{A_i}(x) \underset{(18.360)}{=} \sum_{i=1}^n \alpha_i \cdot \mathcal{X}_{E, A_i \bigcap E}(x)$, proving that $f_{|E}$ has the representation $\sum_{i=1}^n \alpha_i \cdot \mathcal{X}_{E, A_i \bigcap E}$ which means that $f_{|E} \in \mathcal{S}[E, \mathcal{A}_E]$.

Finally we have that

$$\begin{aligned} \int^S f_{|E} d\mu_E &= \sum_{i=1}^n \alpha_i \cdot \mu_E\left(A_i \bigcap E\right) \\ &\underset{18.152}{=} \sum_{i=1}^n \alpha_i \cdot \mu\left(A_i \bigcap E\right) \\ &\leqslant_{18.20\ (2)} \sum_{i=1}^n \alpha_i \cdot \mu(A_i) \\ &= \int^S f d\mu \end{aligned}$$

proving that

$$\int^S f_{|E} d\mu_E \leqslant \int^S f d\mu$$

and the lemma. $\square$

Lemma 18.167. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space, $E \in \mathcal{A}$ and $f \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ then $f_{|E} \in \mathcal{L}[E, \mathcal{A}_E, \mu_E]$*

Proof. Using the definition $\mathcal{L}_+[X, \mathcal{A}, \mu]$ [see Definition 18.126] we have

$$0 \leqslant f \wedge \exists \{f_i\}_{i \in \mathbb{N}} \subseteq \mathcal{S}[X, \mathcal{A}] \text{ such that } \forall n \in \mathbb{N} \quad 0 \leqslant f_n \leqslant f_{n+1}, \sup\left(\left\{\int^S f_n d\mu | n \in \mathbb{N}\right\}\right) < \infty \text{ and } \lim_{n \to \infty} f_n =_{a.e.[\mu]} f \tag{18.361}$$

Further $\forall x\in \operatorname{dom}(f_{|E})=\operatorname{dom}(f)\bigcap E$ we have $f_{|E}(x)=f(x)\in[0,\infty[$ proving that

$$0\leqslant f_{|E} \tag{18.362}$$

Let $x\in E$ then $\forall n\in\mathbb{N}$ we have $0\leqslant f_n(x)=(f_n)_{|E}(x)$ and $(f_n)_{|E}(x)=f_n(x)\leqslant f_{n+1}(x)=(f_{n+1})_{|E}(x)$ proving that

$$\forall n\in\mathbb{N}\text{ we have }0\leqslant(f_n)_{|E}\leqslant(f_{n+1})_{|E} \tag{18.363}$$

Using the previous lemma (see 18.166) we have then

$$\{(f_n)_{|E}\}_{n\in\mathbb{N}}\subseteq\mathcal{S}[E,\mathcal{A}_E] \tag{18.364}$$

Using 18.124 we have that $\operatorname{dom}\left(\lim_{n\to\infty}f_n\right)=\left\{x\in X|\lim_{n\to\infty}f_n(x)\text{ exists}\right\}$, further as $\lim_{n\to\infty}f_n=_{a.e.[\mu]}f$ we have that $A=\left\{x\in\operatorname{dom}\left(\lim_{n\to\infty}f_n\right)\bigcap\operatorname{dom}(f)|f(x)=\lim_{n\to\infty}f_n(x)\right\}$ is μ-conegligible. Further

$$\begin{aligned}
x\in\operatorname{dom}\left(\lim_{n\to\infty}f_n\right)\bigcap E \quad&\Leftrightarrow\quad x\in X\bigcap E\wedge\lim_{n\to\infty}f_n(x)\text{ exists}\\
&\underset{x\in E,E\subseteq X}{\Leftrightarrow}\quad x\in E\wedge\lim_{n\to\infty}(f_n)_{|E}(x)\text{ exists}\\
&\underset{(18.364)\text{ and }18.124\ (1)}{\Leftrightarrow}\quad x\in\operatorname{dom}\left(\lim_{n\to\infty}(f_n)_{|E}\right)
\end{aligned}$$

proving that

$$\operatorname{dom}\left(\lim_{n\to\infty}f_n\right)\bigcap E=\operatorname{dom}\left(\lim_{n\to\infty}(f_n)_{|E}\right) \tag{18.365}$$

So

$$\begin{aligned}
x\in A\bigcap E \quad&\Leftrightarrow\quad x\in\operatorname{dom}\left(\lim_{n\to\infty}f_n\right)\bigcap\operatorname{dom}(f)\bigcap E\wedge\lim_{n\to\infty}f_n(x)\text{ exists}\\
&\underset{(18.365)}{\Leftrightarrow}\quad x\in\operatorname{dom}\left(\lim_{n\to\infty}(f_n)_{|E}\right)\bigcap\operatorname{dom}(f_{|E})\bigcap E\wedge\lim_{n\to\infty}f_n(x)\text{ exists}\\
&\underset{x\in E}{\Leftrightarrow}\quad x\in\operatorname{dom}\left(\lim_{n\to\infty}(f_n)_{|E}\right)\bigcap\operatorname{dom}(f_{|E})\wedge\lim_{n\to\infty}(f_n)_{|E}(x)\text{ exists}\\
&\Leftrightarrow\quad x\in\left\{x\in\operatorname{dom}\left(\lim_{n\to\infty}(f_n)_{|E}\right)\bigcap\operatorname{dom}(f_{|E})|\lim_{n\to\infty}(f_n)_{|E}(x)\text{ exists}\right.
\end{aligned}$$

which proves that $\left\{x\in\operatorname{dom}\left(\lim_{n\to\infty}(f_n)_{|E}\right)\bigcap\operatorname{dom}(f_{|E})|\lim_{n\to\infty}(f_n)_{|E}(x)\text{ exists}\right\}=A\bigcap E$ which is μ_E-conegligible [see 18.153] proving that

$$\lim_{n\to\infty}(f_n)_{|E}=_{a.e.[\mu_E]}f_{|E} \tag{18.366}$$

Further as $\forall I\in\left\{\int^S(f_n)_{|E}d\mu_E|n\in\mathbb{N}\right\}$ we have $\exists n\in\mathbb{N}$ such that $I=\int^S(f_n)_{|E}d\mu_E\leqslant_{18.166}\int^S f_nd\mu\in\left\{\int^S f_nd\mu|n\in\mathbb{N}\right\}$ proving by 2.172 that

$$\sup\left(\left\{\int^S(f_n)_{|E}d\mu_E|n\in\mathbb{N}\right\}\right)\leqslant\sup\left(\left\{\int^S f_nd\mu|n\in\mathbb{N}\right\}\right)<\infty \tag{18.367}$$

Using (18.362),(18.363),(18.364),(18.367) and (18.366) allows to apply the definition of positive integrable functions proving that

$$f_{|E}\in\mathcal{L}_+[E,\mathcal{A}_E,\mu_E]$$

□

Lemma 18.168. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space, $E \in \mathcal{A}$ and $f \in \mathcal{L}[X, \mathcal{A}, \mu]$ then $f_{|E} \in \mathcal{L}[E, \mathcal{A}_E, \mu_E]$*

Proof. As $f \in \mathcal{L}[X, \mathcal{A}, \mu]$ there exists $f_1, f_2 \in \mathcal{L}_+[X, \mathcal{A}, \mu]$ such that $f = f_1 - f_2$. Using the previous lemma [see 18.167] we have that $(f_1)_{|E}, (f_2)_{|E} \in \mathcal{L}_+[E, \mathcal{A}_E, \mu_E]$. Further $\forall x \in \mathrm{dom}(f_{|E}) = \mathrm{dom}(f) \bigcap E$ we have $f_{|E}(x) = f(x) = f_1(x) - f_2(x) = (f_1)_{|E}(x) - (f_2)_{|E}(x)$ proving that $f_{|E} = (f_1)_{|E} - (f_2)_{|E}$. So by the definition of integrable functions [see 18.136] we have that $f_{|E} \in \mathcal{L}[E, \mathcal{A}_E, \mu_E]$. □

Proposition 18.169. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure space $f: X \rightarrow \mathbb{R}$ a partial function then*

1. *If $E \in \mathcal{A}$ and $\mathrm{dom}(f)$ is μ-conegligible then $f_{|E} \in \mathcal{L}[E, \mathcal{A}_E, \mu_E]$ if and only if $f \cdot \mathcal{X}_E \in \mathcal{L}[X, \mathcal{A}, \mu]$. Further if $f_{|E} \in \mathcal{L}[E, \mathcal{A}_E, \mu_E]$ [or equivalent $f \cdot \mathcal{X}_E \in \mathcal{L}[X, \mathcal{A}, \mu]$] we have $\int_E f d\mu = \int f \cdot \mathcal{X}_E d\mu$*
2. *Let $f \in \mathcal{L}[X, \mathcal{A}, \mu]$ and $E \in \mathcal{A}$ then $f_{|E} \in \mathcal{L}[E, \mathcal{A}_E, \mu_E]$, $f \cdot \mathcal{X}_E \in \mathcal{L}[X, \mathcal{A}, \mu]$ and $\int_E f d\mu = \int f \cdot \mathcal{X}_E d\mu$*
3. *If $f \in \mathcal{L}[X, \mathcal{A}, \mu]$ then $0 \leqslant_{a.e.} f$ if and only if $0 \leqslant \int_E f d\mu$ $\forall E \in \mathcal{A}$ [note that because of (2) $\int_E f d\mu$ is defined]*
4. *If $f \in \mathcal{L}[X, \mathcal{A}, \mu]$ then $f =_{a.e.} 0$ if and only if $\int_E f d\mu = 0$ $\forall E \in \mathcal{A}$ [note that because of (2) $\int_E f d\mu$ is defined]*

Proof.

1. Let $x \in \mathrm{dom}(f)$ then we have either

   $x \notin E$. then $x \in X \setminus E$ and $(f_{|E})^E(x) = 0 = f(x) \cdot 0 = f(x) \cdot \mathcal{X}_E(x) = (f \cdot \mathcal{X}_E)(x)$

   $x \in E$. then $x \in \mathrm{dom}(f_{|E})$ so that $(f_{|E})^E(x) = f_{|E}(x) = f(x) \underset{x \in E}{=} f(x) \cdot \mathcal{X}_E(x) = (f \cdot \mathcal{X}_E)(x)$

   which as $\mathrm{dom}(f)$ is μ-conegligible means that

   $$(f_{|E})^E =_{a.e.[\mu]} f \cdot \mathcal{X}_E \tag{18.368}$$

   If now $f_{|E} \in \mathcal{L}[E, \mathcal{A}_E, \mu_E]$ we have by the previous corollary [see 18.164] that $(f_{|E})^E \in \mathcal{L}[X, \mathcal{A}, \mu]$ and applying (18.368) on 18.145 (2) proves then that $f \cdot \mathcal{X}_E \in \mathcal{L}[X, \mathcal{A}.\mu]$. On the other hand if $f \cdot \mathcal{X}_E \in \mathcal{L}[X, \mathcal{A}, \mu]$ we have applying (18.368) on 18.145 (2) that $(f_{|E})^E \in \mathcal{L}[X, \mathcal{A}, \mu]$ which by the previous corollary 18.164 proves that $f_{|E} \in \mathcal{L}[E, \mathcal{A}_E, \mu]$. So we have

   $$f_{|E} \in \mathcal{L}[E, \mathcal{A}_E, \mu_\mu] \Leftrightarrow f \cdot \mathcal{X}_E \in \mathcal{L}[X, \mathcal{A}, \mu]$$

   Further if $f_{|E} \in \mathcal{L}[E, \mathcal{A}_E, \mu_\mu]$ then $\int_E f\, d\mu \underset{\text{def}}{=} \int f_{|E} d\mu_E \underset{18.164}{=} \int (f_{|E})^E d\mu \underset{18.145(2) \text{ and } (18.368)}{=} \int f \cdot \mathcal{X}_E d\mu$ proving

   $$\int_E f d\mu = \int f \cdot \mathcal{X}_E d\mu$$

2. Let $f \in \mathcal{L}[X, \mathcal{A}, \mu]$ and $E \in \mathcal{A}$ then by 18.168 $f_{|E} \in \mathcal{L}[E, \mathcal{A}_E, \mu]$ so that by (1) $f \cdot \mathcal{X}_E \in \mathcal{L}[X, \mathcal{A}, \mu]$ and $\int_E f d\mu = \int f \cdot \mathcal{X}_E d\mu$

3. First as $f \in \mathcal{L}[X, \mathcal{A}, \mu]$ we have by (2) that $f_{|E} \in \mathcal{L}[E, \mathcal{A}_E, \mu_E]$. If $0 \leqslant_{a.e.} f$ then by Algorithm 18.154 we have $\forall E \in \mathcal{A}$ that $0 \leqslant_{a.e.[\mu_E]} f_{|E}$, so that by 18.141 $0 \leqslant \int f_{|E} d\mu_E \underset{\text{def}}{=} \int_E f d\mu$, proving

$$0 \leqslant_{a.e.[\mu]} f \Rightarrow \forall E \in \mathcal{A} \text{ we have } 0 \leqslant \int_E f d\mu \tag{18.369}$$

For the opposite implication, assume that $\forall E \in \mathcal{A}\ 0 \leqslant \int_E f d\mu$. As $f \in \mathcal{L}[X.\mathcal{A}.\mu]$ there exists by 18.143 a μ-conegligible set D such that $f_{|D}$ is $\mathcal{A}$-measurable. As by 18.137 $\text{dom}(f)$ is μ-conegligible there exists by 18.27 a μ-conegligible set $E \in \mathcal{A}$ such that $E \subseteq \text{dom}(f) \bigcap D$. Further as $f_{|E} \underset{E \subseteq D \text{ and } 2.28}{=} (f_{|D})_{|E}$ we have by 18.103 (7) that $f_{|E}$ is $\mathcal{A}$-measurable. Summarizing we have

$$\exists E \in \mathcal{A} \quad \text{such that } E \text{ is } \mu\text{-conegligible, } E \subseteq \text{dom}(f) \text{ and } f_{|E} \text{ is } \mathcal{A}\text{-measurable} \tag{18.370}$$

As $f_{|E}$ is $\mathcal{A}$-measurable we have that $F = \{f_{|E} < 0\} \in \mathcal{A}$ where $\{f_{|E} < 0\}$ Also using 18.27 (1) F is μ_F-conegligible. Further if $x \in F$ then $f_{|F}(x) = f(x) \underset{F \subseteq E}{=} f_{|E}(x) < 0 \Rightarrow f_{|F}(x) \leqslant 0$ proving that $F \subseteq \{x \in \text{dom}(f_{|F}) | f_{|F}(x) \leqslant 0\}$. Using 18.27 (2) it follows that $\{x \in \text{dom}(f_{|F}) | f_{|F}(x) \leqslant 0\}$ is μ_F-conegligible, so that $f_{|F} \leqslant_{a.e.[\mu_F]} 0$. Further we have by assumption that $0 \leqslant \int_F f d\mu = \int f_{|F} d\mu_F$, hence we can apply 18.145 (4) giving $f_{|F} =_{a.e.[\mu_F]} 0$. So by definition $G = \{x \in \text{dom}(f_{|F}) | f(x) = 0\}$ is μ_F-conegligible. As $\forall x \in F$ we have $f(x) < 0$ we must have that $G = \emptyset$. Using the μ_F-Conegligible of G it follows that $\mu(F) \underset{G = \emptyset}{=} \mu(F \setminus G) = \mu_F(F \setminus G) = 0$, so we have

$$F \in \mathcal{A} \text{ and } \mu(F) = 0 \text{ so that } F \text{ is } \mu\text{-negligible} \tag{18.371}$$

Take now $x \in X \setminus \{0 \leqslant f\} = X \setminus \{x \in \text{dom}(f) | 0 \leqslant f(x)\}$ then we must consider the following possible cases

$x \in X \setminus \text{dom}(f)$. then $x \in (X \setminus E) \bigcup (X \setminus \text{dom}(f)) \bigcup F$

$x \in \text{dom}(f)$. then $f(x) < 0$ and we have either

- **$x \in E$.** then $x \in \text{dom}(f) \bigcap E = \text{dom}(f_{|E})$
  - **$x \in F$.** then $x \in (X \setminus E) \bigcup (X \setminus \text{dom}(f)) \bigcup F$
  - **$x \notin F$.** then $0 \leqslant f(x)$ [for if $f(x) < 0$ then $f_{|E}(x) < 0 \Rightarrow x \in \{f_{|E} < 0\} = F$] contradicting $f(x) < 0$ so this case will never occur.
- **$x \notin E$.** then $x \in (X \setminus E) \bigcup (X \setminus \text{dom}(f)) \bigcup F$

so we have

$$X \setminus \{0 \leqslant f\} = X \setminus \{x \in \text{dom}(f) | 0 \leqslant f(x)\} \subseteq (X \setminus E) \bigcup (X \setminus \text{dom}(f)) \bigcup F \tag{18.372}$$

Now as $f \in \mathcal{L}[X, \mathcal{A}, \mu]$ we have by 18.137 that $\operatorname{dom}(f)$ is μ-conegligible so that by definition $X \setminus \operatorname{dom}(f)$ is μ-negligible, further as E is μ-conegligible we have by definition that $X \setminus E$ is μ-negligible. Using this and (18.371) together with 18.24 (3) proves that $(X \setminus E) \bigcup (X \setminus \operatorname{dom}(f)) \bigcup F$ is μ-negligible. Applying then 18.24 (2) proves that $X \setminus \{x \in \operatorname{dom}(f) | 0 \leqslant f(x)\}$ is μ-negligible hence $\{x \in \operatorname{dom}(f) | 0 \leqslant f(x)\}$ is μ-conegligible proving that $0 \leqslant_{a.e.[\mu]} f$. As $E \in \mathcal{A}$ was chosen arbitrary we have

$$\text{If} \, \forall E \in \mathcal{A} \, \text{we have} \, 0 \leqslant \int_E f d\mu \, \text{then} \, 0 \leqslant_{a.e.[\mu]} f$$

which together with (18.369) proves that

$$0 \leqslant_{a.e.[\mu]} f \Leftrightarrow \forall E \in \mathcal{A} \, \text{we have} \, 0 \geqslant \int_E f d\mu \tag{18.373}$$

4. If $f =_{a.e.[\mu]} 0$ then by 18.30 $0 \leqslant_{a.e.[\mu]} f \leqslant_{a.e.[\mu]} 0$ and $f \leqslant_{a.e.[\mu]} 0 \underset{18.30}{\Rightarrow} 0 \leqslant_{a.e.[\mu]} -f$ so that by (3) we have that $\forall E \in \mathcal{A}$ $0 \leqslant \int_E f d\mu \wedge \int_E f d\mu \leqslant 0 \Rightarrow 0 \leqslant -\int_E f d\mu = \int_E -f d\mu$ proving that $0 = \int_E f d\mu$. For the opposite implication assume that $0 = \int_E f d\mu$ then $0 \leqslant \int_E f d\mu$ and $\int_E f d\mu \leqslant 0 \Rightarrow 0 \leqslant -\int_E f d\mu = \int_E -f d\mu$ so that by (3) we have $0 \leqslant_{a.e.[\mu]} f$ and $0 \leqslant_{a.e.[\mu]} -f \underset{18.30}{\Rightarrow} f \leqslant_{a.e.[\mu]} 0$. Using then 18.30 it follows that $f =_{a,e,[\mu]} 0$. □

Corollary 18.170. *Let $\langle X, \mathcal{A}, \mu \rangle$ be a measure space and $f, g \in \mathcal{L}[X, \mathcal{A}, \mu]$ then we have*

1. *If $\forall E \in \mathcal{A}$ $\int_E f d\mu \geqslant \int_E g d\mu$ then $f \geqslant_{a.e.} g$*
2. *If $\forall E \in \mathcal{A}$ $\int_E f d\mu = \int_E g d\mu$ then $f =_{a.e.} g$*

Proof. *Take $h = f - g$ then we have if $E \in \mathcal{A}$ by 18.141 that $h_{|E} = f_{|E} - g_{|E} \in \mathcal{L}[E, \mathcal{A}_E, \mu_E]$ and $\int_E h d\mu = \int h_{|E} d\mu_E = \int f_{|E} d\mu_E - \int g_{|E} d\mu_E = \int_E f d\mu - \int g d\mu$. So*

1. *If $\forall E \in \mathcal{A}$ we have that $\int_E f d\mu \geqslant \int_E g d\mu$ then $0 \leqslant \int_E h d\mu$ so by 18.169 we have that $0 \leqslant_{a.e.} h$ so $H = \{x \in \operatorname{dom}(h) = \operatorname{dom}(f) \bigcap \operatorname{dom}(g) | 0 \leqslant h(x)\}$ is μ-conegligible. If $x \in H$ then $0 \leqslant h(x) \Rightarrow 0 \leqslant f(x) - g(x) \Rightarrow g(x) \leqslant f(x)$ proving that $f \leqslant_{a.e.} g$*
2. *If $\forall E \in \mathcal{A}$ we have that $\int_E f d\mu = \int_E g d\mu$ then $0 = \int_E h d\mu$ so by 18.169 we have that $0 =_{a.e.} h$ so $H = \{x \in \operatorname{dom}(h) = \operatorname{dom}(f) \bigcap \operatorname{dom}(g) | 0 = h(x)\}$ is μ-conegligible. If $x \in H$ then $0 = h(x) \Rightarrow 0 = f(x) - g(x) \Rightarrow g(x) = f(x)$ proving that $f \leqslant_{a.e.} g$* □

Proposition 18.171. *Let $\langle X, \mathcal{A}, \mu \rangle$ be a measure space, $E \in \mathcal{A}$ such that E is μ-conegligible then we have $f \in \mathcal{L}[X, \mathcal{A}, \mu]$ if and only if $f_{|E} \in \mathcal{L}[X, \mathcal{A}, \mu]$ and if $f \in \mathcal{L}[X, \mathcal{A}, \mu]$ [or equivalently $f_{|E} \in \mathcal{L}[X, \mathcal{A}, \mu]$] then $\int_E f d\mu = \int f d\mu$.*

Proof. First note that by 18.168 we have

$$\forall E \in \mathcal{A} \, \text{we have} \, f \in \mathcal{L}[X, \mathcal{A}, \mu] \Rightarrow f_{|E} \in \mathcal{L}[X, \mathcal{A}, \mu] \tag{18.374}$$

Further take E and assume that $f_{|E} \in \mathcal{L}[X, \mathcal{A}, \mu]$ then by 18.164 we have that

$$(f_{|E})^E \in \mathcal{L}[X, \mathcal{A}, \mu] \text{ and } \int f_{|E} d\mu_E = \int (f_{|E})^E d\mu \tag{18.375}$$

Now $\mathrm{dom}(f_{|E})^E = \mathrm{dom}(f_{|E}) \bigcup (X \setminus E) = (\mathrm{dom}(f) \bigcap E) \bigcup (X \setminus E)$ and as $(f_{|E})^E \in \mathcal{L}[X\mathcal{A}, \mu]$ we have by 18.137 that $(\mathrm{dom}(f) \bigcap E) \bigcup (X \setminus E)$ is μ-conegligible. As E is μ-conegligible we have that $X \setminus E$ is μ-negligible and we can then use 18.27 (7) giving that

$$\mathrm{dom}(f_{|E}) \text{ is } \mu\text{-conegligible} \tag{18.376}$$

Further if $x \in \mathrm{dom}(f_{|E}) = \mathrm{dom}(f) \bigcap E$ then $(f_{|E})^E(x) \underset{x \in \mathrm{dom}(f_{|E})}{=} f_{|E}(x) = f(x)$ proving by (18.376) that

$$f_{|E} =_{a.e.[\mu]} f \tag{18.377}$$

Using 18.145 together with (18.375) and the above proves then that $f \in \mathcal{L}[X, \mathcal{A}, \mu]$ and $\int f d\mu = \int (f_{|E})^E d\mu \underset{(18.375)}{=} \int f_{|E} d\mu \underset{\text{def}}{=} \int_E f d\mu$. Using this result together with (18.374) proves that if $E \in \mathcal{A}$ such that E is μ-conegligible then

$$f \in \mathcal{L}[X, \mathcal{A}, \mu] \Leftrightarrow f_{|E} \in \mathcal{L}[E, \boldsymbol{A}_E, \mu_E] \text{ and in both cases } \int_E f d\mu = \int f d\mu \qquad \square$$

18.5 Outer measure induced by a measure

Given a measure space we can use lemma 18.35 to define a outer measure based on a measure space, to formalize this we state the following definition.

Definition 18.172. *Let $\langle X, \mathcal{A}, \mu \rangle$ be a measure space then $\mu^*: \mathcal{P}(x) \rightarrow [0, \infty]$ is defined by $\mu^*(A) = \inf(\{\mu(E) | E \in \mathcal{A} \wedge A \subseteq E\})$. μ^* is called the **outer measure induced by** μ.*

In lemma 18.35 we have already proved that μ^* is indeed a outer measure on X, in the next proposition we sum up the other properties of the induced outer measure,

Proposition 18.173. *Let $\langle X, \mathcal{A}, \mu \rangle$ be a measure space then we have the following properties for $\mu^*: \mathcal{P}(X) \rightarrow [0, \infty]$*

1. *$\forall A \subseteq X$ there exists a $E \in \mathcal{A}$ such that $A \subseteq E$ and $\mu^*(A) = \mu(E)$*
2. *μ^* is a outer measure (see 18.32)*
3. *$\forall E \in \mathcal{A}$ we have $\mu^*(E) = \mu(E)$*
4. *$A \subseteq X$ is μ-negligible if and only if $\mu^*(A) = 0$*
5. *If $\{A_i\}_{i \in \mathbb{N}}$ such that $A_i \subseteq A_{i+1}$ we have $\mu^*(\bigcup_{i \in \mathbb{N}} A_i) = \sup(\{\mu^*(A_i) | i \in \mathbb{N}\}) = \lim_{n \rightarrow \infty} \mu^*(A_n)$*
6. *$\forall A \subseteq X$ and $\forall E \in \mathcal{A}$ we have $\mu^*(A) = \mu^*(A \bigcap E) + \mu^*(A \setminus E)$*

Proof.

1. This follows from 18.35 (1)
2. This follows from 18.35 (2)
3. This follows from 18.35 (3)
4. Assume that A is μ-negligible then there exists a $E\in\mathcal{A}$ with $\mu(E)=0$ so that $\mu^*(A)=\inf(\{\mu(E)|E\in\mathcal{A}\wedge E\subseteq E\})\leqslant\mu(E)=0$ proving that $\mu^*(A)=0$. On the other hand if $\mu^*(A)=0$ then by (1) there exists a $E\in\mathcal{A}$ such that $A\subseteq E$ and $\mu(E)=\mu^*(A)=0$ proving that A is μ-negligible.
5. As a outer measure is increasing we have by definition [see 18.32 (2)] that $\forall i\in\mathbb{N}\ \mu^*(A_i)\leqslant\mu^*(A_{i+1})$. Now $\forall i\in\mathbb{N}$ there exists a $E_i\in\mathcal{A}$ such that $A_i\subseteq E_i$ and $\mu^*(A)=\mu(E)$. For $n\in\mathbb{N}$ define $F_n=\bigcap_{i\in\{n,\ldots,\infty\}}E_i$ then $F_n=\bigcap_{i\in\{n,\ldots,\infty\}}E_i\subseteq\bigcap_{i\in\{n+1,\ldots,\infty\}}E_i$. Using 18.8 it follows that $\{F_n\}_{n\in\mathbb{N}}\subseteq\mathcal{A}$ hence we can use 18.20 (5) we have that

$$\mu\left(\bigcup_{n\in\mathbb{N}}F_n\right)=\lim_{n\to\infty}\mu(F_n)=\sup(\{\mu(F_n)|n\in\mathbb{N}\})\tag{18.378}$$

Let $n\in\mathbb{N}$ and assume that $x\in A_n$ then $\forall i\in\{n,\ldots,\infty\}$ we have $x\in A_n\subseteq A_i\subseteq E_i$ so that $x\in\bigcap_{i\in\{n,\ldots,\infty\}}E_i=F_n\subseteq E_n$ so that $A_n\subseteq F_n\subseteq E_n$ so that $\mu^*(A_n)\leqslant_{18.32\ (2)}\mu^*(F_n)\underset{(3)}{=}\mu(F_n)\leqslant\mu(E_n)=\mu^*(E_n)$ proving that

$$\forall n\in\mathbb{N}\,\mu^*(A_n)=\mu(F_n)\tag{18.379}$$

Using this on (18.378) proves

$$\mu\left(\bigcup_{n\in\mathbb{N}}F_n\right)=\lim_{n\to\infty}\mu^*(A_n)=\sup(\{\mu^*(A_n)|n\in\mathbb{N}\})\tag{18.380}$$

Also from $A_n\subseteq F_n$ we have that $\bigcup_{n\in\mathbb{N}}A_n\subseteq\bigcup_{n\in\mathbb{N}}F_n\in\mathcal{A}$ hence $\mu^*(\bigcup_{n\in\mathbb{N}}A_n)\leqslant\mu^*(\bigcup_{n\in\mathbb{N}}F_n)\underset{(3)}{=}\mu(\bigcup_{n\in\mathbb{N}}F_n)\underset{(18.380)}{=}\lim_{n\to\infty}\mu(F_n)\underset{(18.379)}{=}\lim_{n\to\infty}\mu^*(A_n)=\sup(\{\mu^*(A_n)|n\in\mathbb{N}\})$. As also $\forall n\in\mathbb{N}\ A_n\subseteq\bigcup_{i\in\mathbb{N}}A_i$ we have that $\sup(\{\mu^*(A_n)|n\in\mathbb{N}\})\leqslant_{18.32\ (2)}\mu^*(\bigcup_{n\in\mathbb{N}}A_n)$ so that we have

$$\mu^*\left(\bigcup_{\text{in}\in\mathbb{N}}A_n\right)=\lim_{n\to\infty}\mu^*(A_n)=\sup(\{\mu^*(A_n)|n\in\mathbb{N}\})$$

6. As by (3) μ^* is a outer measure we have as $A=(A\bigcap E)\bigcup(A\setminus E)$ by 18.32 (3) that $\mu^*(A)\leqslant\mu^*(A\bigcap E)+\mu^*(A\setminus E)$. For the opposite inclusion note that by (1) there exists a $F\in\mathcal{A}$ such that $A\subseteq F$ and $\mu^*(A)=\mu(F)$. Hence $A\bigcap E\subseteq F\bigcap E\in\mathcal{A}$ and $A\setminus E\subseteq F\setminus E\in\mathcal{A}$ [see 18.8] so that $\mu^*(A\bigcap E)+\mu^*(E\setminus F)\leqslant_{18.32\ (2)}\mu^*(F\bigcap E)+\mu^*(F\setminus E)\underset{(3)}{=}\mu(F\bigcap E)+\mu(F\setminus E)=\mu(F)=\mu^*(A)$. So we have that

$$\forall E\in\mathcal{A}\,\text{we have}\,\mu^*(A)=\mu^*\left(A\bigcap E\right)+\mu^*(A\setminus E)$$

□

We have seen how given a outer measure, we can use the Carathéodory's method to generate a measure based on the outer measure. It is in general however not true that the outer measure induced by this measure is the original outer measure. However in case of the Lebesgue measure on $\mathbb{R}^n$ this is true.

Proposition 18.174. *Let $n \in \mathbb{N}$ then $(\lambda^n)^* = \varphi^n$ [see 18.72] where $\varphi^n(A) = \inf(\{\sum_{i=1}^{\infty} v^n(R_i) | \{R_i\}_{i\in\mathbb{N}} \subseteq \mathcal{R}^n \text{ where } A \subseteq \bigcup_{i\in\mathbb{N}} R_i\})$ is the outer measure on $\mathbb{R}^n$ and $\lambda^n = (\varphi^n)_{|\mathcal{L}^n}$ is the Lebesgue measure in the measurable space $\langle \mathbb{R}^n, \mathcal{L}^n\rangle$*

Proof. Let $A \subseteq \mathbb{R}^n$. If $E \in \mathcal{L}^n$ such that $A \subseteq E$ then as φ^n is a outer measure [see 18.72] we have that $\varphi^n(A) \leqslant \varphi^n(E) = (\varphi^n)_{|\mathcal{L}^n}(E) = \lambda^n(E)$. So $\varphi^n(A) \leqslant \inf(\{\lambda^n(E) | E \in \mathcal{L}^n \wedge A \subseteq E\}) \underset{\text{def}}{=} (\lambda^n)^*(A)$ proving that

$$\varphi^n(A) \leqslant (\lambda^n)^*(A) \tag{18.381}$$

For the opposite inequality let $\varepsilon > 0$ then as $\varphi^n(A) = \inf(\{\sum_{i=1}^{\infty} v^n(R_i) | \{R_i\}_{i\in\mathbb{N}} \text{ where } A \subseteq \bigcup_{i\in\mathbb{N}} R_i\})$ we have that there exists a $\{R_i\}_{i\in\mathbb{N}} \subseteq \mathcal{R}^n$ with $A \subseteq \bigcup_{i\in\mathbb{N}} R_i$ such that $\sum_{i=1}^{\infty} v^n(R_i) < \varphi^n(A) + \varepsilon$. As $A \subseteq \bigcup_{i\in\mathbb{N}} R_i \in \mathcal{L}^n$ we have that $(\lambda^n)^*(A) \leqslant (\lambda^n)^*(\bigcup_{i\in\mathbb{N}} R_i) \underset{18.173\ (3)}{=} \lambda^n(\bigcup_{i\in\mathbb{N}} R_i) \leqslant_{18.20} \sum_{i=1}^{\infty} \lambda^n(R_i) \underset{18.71}{=} \sum_{i=1}^{\infty} v^n(R_i) < \varphi^n(A) + \varepsilon$. As ε is chosen arbitrary we can use 9.56 giving $(\lambda^n)^*(A) \leqslant \varphi^n(A)$ which together with (18.381) proves $(\lambda^n)^*(A) = \varphi^n(A)$. Hence

$$(\lambda^n)^* = \varphi^n$$

□

Definition 18.175. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure and $A \subseteq X$ then a* ***measurable envelop of*** *A is a set $E \in \mathcal{A}$ such that $A \subseteq E$ and $\forall F \in \mathcal{A}$ $\mu(E \bigcap F) = \mu^*(F \bigcap A)$*

Theorem 18.176. *Let $\langle X, \mathcal{A}, \mu\rangle$ be a measure and $A \subseteq \mathcal{A}$ then we have*

1. *If $E \in \mathcal{A}$ such that $A \subseteq E$ then E is a measurable envelop of A if and only if $\forall F \in \mathcal{A}$ with $F \subseteq E \setminus A$ we have $\mu(F) = 0$*
2. *If $E \in \mathcal{A}$ such that $A \subseteq E$ and $\mu(E) < \infty$ then E is a measurable envelop of A if and only if $\mu(E) = \mu^*(A)$*
3. *If E is a measurable envelop of A and $H \in \mathcal{A}$ then $E \bigcap H$ is a measurable envelop of $A \bigcap H$*
4. *Let $\{A_n\}_{n\in\mathbb{N}}$ be a sequence of subsets of X such that $\forall n \in \mathbb{N}$ A_n has a measurable envelop E_n then $\bigcup_{n\in\mathbb{N}} A_n$ is a measurable envelop of $\bigcup_{n\in\mathbb{N}} A_n$*
5. *Let $A \subseteq X$ such that there exists a $\{F_n\}_{n\in\mathbb{N}} \subseteq \mathcal{A}$ with $\mu(F_n) < \infty$ $\forall n \in \mathbb{N}$ such that $A \subseteq \bigcup_{n\in\mathbb{N}} F_n$ then A has a measurable envelop.*

Proof.

1. Let E be a measurable envelop of A then if $F \in \mathcal{A}$ such that $F \subseteq E \setminus A$ then by definition we have $\mu(E \bigcap F) = \mu^*(A \bigcap F)$. As $F \subseteq E \setminus A \subseteq E$ we have $F = F \bigcap E$, further $A \bigcap F = \emptyset$ so that $\mu^*(A \bigcap F) = \mu^*(\emptyset) \underset{18.32}{=} 0$ so we have $\mu(F) = 0$. Hence we have

   If E is a measurable envelop of A then $\forall F \in \mathcal{A}$ with $F \subseteq E \setminus A$ we have $\mu(F) = 0$

For the opposite implication let

$$\forall F \in \mathcal{A} \text{ with } F \subseteq E \setminus A \text{ we have } \mu(F) = \emptyset \tag{18.382}$$

and assume that E is not a measurable envelop of A. As E is not a measurable envelop of A there exists a $H \in \mathcal{A}$ such that $A \subseteq H$ $\mu(E \bigcap H) \neq \mu^*(A \bigcap H)$. As $\mu^*(A \bigcap H) \leqslant_{A \subseteq E} \mu^*(E \bigcap H) \underset{18.8 \text{ and} 18.173}{=} \mu(E \bigcap H)$ it follows that

$$\mu^*\left(A \bigcap H\right) < \mu\left(E \bigcap H\right) \tag{18.383}$$

Using 18.173 (1) there exists a $G \in \mathcal{A}$ such that $A \bigcap H \subseteq G$ and $\mu^*(A \bigcap H) = \mu(G)$, by (18.383) it follows then that

$$\mu(G) < \mu\left(E \bigcap H\right) \tag{18.384}$$

Take $F = (E \bigcap H) \setminus G \in_{18.8} \mathcal{A}$ then $F \subseteq E$ and using (18.384) and 18.20 (7) we have that $0 < \mu(F)$. To summarize we have

$$F \in \mathcal{A}, F \subseteq E \text{ and } 0 < \mu(F) \tag{18.385}$$

Further $F \bigcap A = ((E \bigcap H) \setminus G) \bigcap A \underset{1.31}{=} (A \bigcap E \bigcap H) \setminus G \underset{A \subseteq E}{=} (A \bigcap H) \setminus G \underset{A \bigcap H \subseteq G}{=} \emptyset$ so that

$$F \subseteq E \setminus A \tag{18.386}$$

Finally (18.385) and (18.386) contradicts with (18.382) so the assumption that E is not a measurable envelop of A is false proving that

If $\forall F \in \mathcal{A}$ with $F \subseteq E \setminus A$ we have $\mu(F) = 0$ then E is a measurable envelop of A

2. Let E be a measurable envelop of A then as $E \in \mathcal{A}$ we have that $\mu^*(A) \underset{A \subseteq E}{=} \mu^*(A \bigcap E) = \mu(E \bigcap E) = \mu(E)$ so that we have

   If E is a measurable envelop of A then $\mu^*(A) = \mu(E)$

   Fort the opposite implication assume that $\mu^*(A) = \mu(E)$ and take a $F \in \mathcal{A}$ such that $F \subseteq E \setminus A$. If $x \in A$ then as $A \subseteq E$ we have that $x \in E$ and $x \notin F$ [otherwise $x \in E \setminus A \Rightarrow x \notin A$ a contradicting $x \in A$] proving that $A \subseteq E \setminus F \in_{18.8} \mathcal{A}$. Hence $\mu(E) = \mu^*(A) \leqslant \mu^*(E \setminus F) \underset{18.173(3)}{=} \mu(E \setminus F)$ and as $\mu(E \setminus F) \leqslant \mu(E)$ [as $E \setminus F \subseteq E$] we have that $\mu(E) = \mu(E \setminus F)$. Now as $E = (E \setminus F) \bigsqcup (E \bigcap F) \underset{F \subseteq E}{=} (E \setminus F) \bigsqcup F$ we have $\mu(E) = \mu(E \setminus F) + \mu(F) = \mu(E) + \mu(F)$ which as $\mu(E) < \infty$ proves that $\mu(F) = 0$. Using (1) it follows then that E is a measurable envelop of A. Hence

   If $\mu^*(A) = \mu(E)$ then E is a measurable envelop of A

3. Let $F \in \mathcal{A}$ such that $F \subseteq (E \bigcap H) \setminus A$ then $F \subseteq E \setminus A$ so by (1) we have $\mu(F) = 0$. As F was chosen arbitrary it follows by (1) again that E is a measurable envelop of A.

4. Take $A = \bigcup_{n \in \mathbb{N}} A_n$ and take $E = \bigcup_{n \in \mathbb{N}} E_n$ then as $\forall n \in \mathbb{N}\, A_n \subseteq E_n$ it follows that $A_n \subseteq A \subseteq \bigcup_{n \in \mathbb{N}} E_n$. Let $F \in \mathcal{A}$ such that $F \subseteq (E \setminus A)$ then $\forall n \in \mathbb{N}$ if $x \in F \bigcap E_n$ we have $x \in E_n$ and $x \notin A \Rightarrow x \notin A_n$ [as $A_n \subseteq A$] so that $F \bigcap E_n \subseteq E_n \setminus A_n$, as by 18.8 $F \bigcap E_n \in \mathcal{A}$ we have by (1) that $\mu(F \bigcap E_n) = 0$. Hence $\mu(F) \underset{F \subseteq E}{=} \mu(E \bigcap F) = \mu(\bigcup_{n \in \mathbb{N}} (F \bigcap E_n)) \leqslant_{18.20} \sum \mu(F \bigcap E_n) = 0$. Applying then (1) again proves that E is a measurable envelop of A.
5. Let $\{F_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A}$ with $\forall n \in \mathbb{N}\ \mu(F_n) < \infty$ be such that $A \subseteq \bigcup_{n \in \mathbb{N}} F_n$. Given $n \in \mathbb{N}$ use 18.173 (1) to find a $E_n \in \mathcal{A}$ such that $A \bigcap F_n \subseteq E_n$ and $\mu(E_n) = \mu^*(A \bigcap F_n) \leqslant \mu^*(F_n) \underset{18.173}{=} \mu(F_n) < \infty$, hence using (2) E_n is a measurable envelop of $A \bigcap F_n$. Using (4) we have that $\bigcup_{n \in \mathbb{N}} E_n$ is a measurable envelop of $\bigcup_{n \in \mathbb{N}} (A \bigcap F_n) = A \bigcap (\bigcup_{n \in \mathbb{N}} F_n) = A$ proving that A has a measurable envelop. □

Corollary 18.177. *Let $n \in \mathbb{N}$ and $\langle \mathbb{R}^n, \mathcal{L}^n, \lambda^n \rangle$ be the Lebesgue measure space defined in 18.72 then every $A \subseteq \mathbb{R}^n$ has a measurable envelop.*

Proof. Let $x \in \mathbb{R}^n$ then using a consequence of the Archimedean property of the real numbers (see 9.55) there exists $\forall i \in \{1, \ldots, n\}$ a $n_i \in \mathbb{N}$, such that $|x_i| < n_i \Rightarrow x_i, -x_i < n_i \Rightarrow x_i \in [-n_i, n_i[$, proving that $x \in \prod_{i \in \{1, \ldots, n\}} [-n_i, n_i[$. Take $n_x = \max(\{n_i | i \in \{1, \ldots, n\}\}) \in \mathbb{N}$ then $x \in \prod_{i \in \{1, \ldots, n\}} [-n_x, n_x[$. Hence if we take $\{R_j\}_{j \in \mathbb{N}}$ by $R_j = \prod_{i \in \{1, \ldots, n\}} [-j, j[\subseteq \mathcal{R}^n$ we have $\mathbb{R}^n \subseteq \bigcup_{n \in \mathbb{N}} [-j, j[$. So if $A \subseteq \mathbb{R}^n$ then $A \subseteq \bigcup_{n \in \mathbb{N}} R_n$ and as $\lambda^n(\prod_{i \in \{1, \ldots, n\}} [-j, j[) = \varphi^n(\prod_{i \in \{1, \ldots, n\}} [-j, j[) \underset{18.71}{=} v^n(\prod_{i \in \{1, \ldots, n\}} [-j, j[) = (2 \cdot j)^n < \infty$ we have by the previous theorem (see 18.176 (5)) that A has a measurable envelop. □

Definition 18.178. *Let $\langle X, \mathcal{A}, \mu \rangle$ be a measure space then $A \subseteq X$ is of **full outer measure** or **thick** if X is a measurable envelop of A.*

Using 18.176 together with the above definition we have

Proposition 18.179. *Let $\langle X, \mathcal{A}, \mu \rangle$ be a measure space then for $A \subseteq X$ we have*

1. *A is of full outer measure if and only if $\forall F \in \mathcal{A}$ we have $\mu^*(F \bigcap A) = \mu(f)$*
2. *A is of full outer measure if and only if $\forall F \in \mathcal{A}$ with $F \subseteq X \setminus E$ we have $\mu(F) = 0$*
3. *If $\mu(X) < \infty$ then A is af full outer measure if and only if $\forall F \in \mathcal{A}$ we have $\mu^*(F \bigcap A) = \mu(F)$*

Proof.

1. As $\forall F \in \mathcal{A}$ we have $F \subseteq X$ so that $\mu(F \bigcap X) = \mu(F)$ we have by definition that A is of full outer measure if and only if $\forall \in \mathcal{A}$ we have $\mu^*(A \bigcap F) = \mu(F)$
2. This follows from the fact that $A \subseteq X$ and 18.176
3. This follows from the fact that $A \subseteq X$ and 18.176 □

Further we have from the fact that $\mathbb{R}^n \in \mathcal{L}^n$ and 18.177that

Proposition 18.180. *Let $n \in \mathbb{N}$ and $\langle \mathbb{R}^n, \mathcal{L}^n, \lambda^n \rangle$ be the Lebesgue measure space defined in 18.72 then every $A \subseteq \mathbb{R}^n$ is of full measure.*

18.6 d

Index